

AD-A184 164

OHIO STATE UNIV COLUMBUS DEPT OF GEQUETIC SCIENCE
THE USE OF FINITE ELEMENTS IN PHYSICAL GEODESY (U)
APR 81 P WEISS

P/O 8/5

UNCLASSIFIED

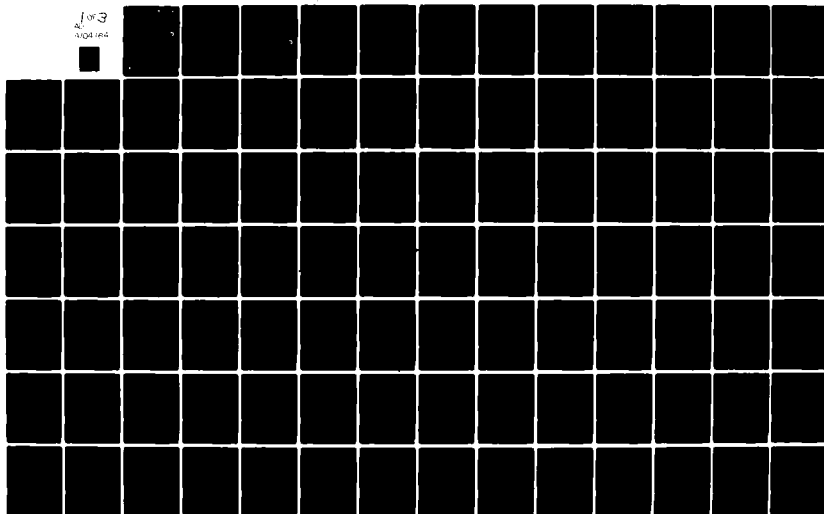
DOS-313

AFGL-TR-81-0114

F19628-79-C-0075

NL

103
ADOLPH



1
LEVEL

12

AFGL-TR-81-0114

AD A104164

THE USE OF FINITE ELEMENTS IN PHYSICAL GEODESY

PETER MEISSL

DEPARTMENT OF GEODETIC SCIENCE
THE OHIO STATE UNIVERSITY
RESEARCH FOUNDATION
COLUMBUS, OHIO 43210

APRIL 1981

SCIENTIFIC REPORT NO. 7

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AIR FORCE GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATE AIR FORCE
HANSCOM AFB, MASSACHUSETTS 01731

DTIC
ELECTE
SEP 14 1981
S D H

FILE COPY

81 9 14 070

Qualified requestors may obtain additional copies from the Defense Documentation Center. All others should apply to the National Technical Information Service.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFGL-TR-81-0114	2. GOVT ACCESSION NO. AD-A104164	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) THE USE OF FINITE ELEMENTS IN PHYSICAL GEODESY		5. TYPE OF REPORT & PERIOD COVERED Scientific Report No. 7
7. AUTHOR(s) PETER MEISSL		6. PERFORMING ORG. REPORT NUMBER Report No. 313
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Geodetic Science The Ohio State University Columbus, Ohio 43210		8. CONTRACT OR GRANT NUMBER(s) F19628-79-C-0075
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Geophysics Laboratory Hanscom AFB, Massachusetts 01731 Contract Monitor - Bela Szabo - L.W.		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62101F 760003A1
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE April 1981
		13. NUMBER OF PAGES Pages 209
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Physical geodesy, Finite elements, Spherical harmonics, Surface layer, Splines, Stokes formula, Vening Meinesz formula, Multipoles, Free boundary value problem, Hermite cubics, Computational efficiency, and Comparison of methods.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Currently used methods of computational physical geodesy are compared with respect to their efficiency during production runs on a computer. These methods include: (1) Least Squares adjustment with respect to spherical harmonics; (2) Surface layers, buried masses and related methods; (3) Least squares collocation; (4) Representation of the potential by spline functions; (5) Explicit integral formulas. As an alternative, the feasibility of applying the finite element method to the fundamental problems of physical geodesy is investigated. The methods listed under (1)-(4) can be dramatically		

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

speeded up if the distribution of data and weights satisfies certain symmetry-requirements which are rather stringent. Method (5) relies altogether on a special type and distribution of data. In the absence of data homogeneity and regularity, the finite element method is asymptotically superior with respect to computational efficiency. Let N denote the number of parameters necessary to describe the variation of the potential on the reference surface. The computational effort associated with methods (1)-(4) grows proportional to N^3 . That one resulting from finite elements grows proportional to N^2 . The constants of proportionality are, however, unfavorable for the finite element method. Hence its superiority comes through only for large values of N , which, in case of a global solution, corresponds to data averaged over $2^0 \times 2^0$ blocks.

FOREWORD

This report has been prepared by Peter Meissl, Professor for Theoretical Geodesy at the Technical University in Graz. Part of the work, in particular that one involving a large computer, was carried out during a 3-month research period at the Department of Geodetic Science, at The Ohio State University. The work was done under Air Force Contract No. F19628-79-C-0075, The Ohio State University Research Foundation, Project No. 711715, Project Supervisor, Urho A. Uotila, Professor, Department of Geodetic Science. The contract covering this research is administered by the Air Force Geophysics Laboratory (AFGL), Hanscom Air Force Base, Massachusetts, with Bela Szabo, Project Scientist.

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Avail and/or	
Dist	Special
A	

ACKNOWLEDGEMENTS

The author wishes to thank Professor U.A. Uotila for valuable advice. Also the discussions with O. Colombo are gratefully acknowledged. Very efficient help in problems of computer-job-control, interactive use of computers, and telecommunication, was received from Lenny A. Krieg. My co-worker B. Hofmann-Wellenhof shared with me the workload of compiling the final report. The manuscript was assembled and plotted with the help of a self-made text-editing system. The system is being developed by my co-workers N. Bartelme and H. Zach. Unfortunately the programs supporting indexing, Greek lettering, and mathematical notation became available at a time when most of the report was already processed. Chapter 7 gives a sample of the full capabilities of the system.

Contents

1.	Introduction and outline of results	1
2.	Review of various methods	6
2.1.	Spherical harmonics	7
2.2.	Surface layer representation and related methods	9
2.3.	Least squares collocation	10
2.3.1.	Krarup's proposal	10
2.3.2.	Least squares collocation using unknown parameters	12
2.4.	Finite elements	13
2.5.	Spline functions	16
2.6.	Approximate explicit Green's functions	19
2.7.	Exploiting rotational or translational symmetries	20
2.7.1.	Invariance of normal equations under a group of transformations	20
2.7.2.	Outline for the case of rotational symmetry around an axis	22
2.7.3.	The effort of the TASC group	23
2.7.4.	Rauhala's array algebra	29
3.	Outline of the finite element approach	33
3.1.	Hermite-cubic representation of the field	33
3.1.1.	The four basis functions for the unit interval	33
3.1.2.	Interval of length h	34
3.1.3.	Bicubic polynomial in a rectangle with sides a, b	35
3.1.4.	Tricubic polynomials in a box with sides a, b, c	36
3.1.5.	C^1 continuity across element boundaries	36
3.1.6.	Compound elements	38
3.2.	Shape functions	40
3.3.	Contribution of the field to the normal equations	41
3.3.1.	Reasons for excluding the Ritz method	41
3.3.2.	The least squares contribution of the field	45
3.4.	Detailed instructions for computer implementation	48
3.4.1.	The partial normals of a simple quad	50
3.4.2.	Partial normals of a compound element	58
3.5.	More general basis functions	63
3.6.	Elements extending to infinity in one direction	67
3.7.	The outer zone	71

3.8.	Local data deficiencies	78
4.	Estimation of computation time	81
4.1.	Nested dissection and Helmert blocking	81
4.2.	Global solution	87
4.3.	The remote zone effect	101
4.4.	Detailed solution in a strip	101
4.5.	Rectangular region	107
4.6.	Further effort to cut down the computation time	111
5.	A proposal for a numerical solution of the free boundary value problem	117
5.1.	Isoparametric elements	117
5.2.	Approaching the free boundary value problem of physical geodesy	125
6.	Computer experiments for 2-dimensional problems	129
6.1.	Purpose and scope	129
6.2.	Parameters distinguishing the experiments	132
6.3.	Detailed results for two experiments	138
6.3.1.	An experiment using comparatively large elements	138
6.3.2.	An experiment using comparatively small elements	165
6.4.	Summary of other experiments	177
6.4.1.	Experiments using bicubics	177
6.4.2.	Experiments using biquintics	179
6.4.3.	Discussion	181
6.4.4.	A word on the Computer programs	182
6.4.5.	Further desirable experiments	182
7.	A proposed hybrid method	184
7.1.	Revision of the surface layer method	184
7.2.	Multipole layer	184
7.3.	Multipole layers on the sphere	186
	Appendix A. Computational effort associated with the partial reduction of a profiled system of normal equations	191
	References	194

1. Introduction and outline of results.

This report pursues two goals. These are

- (1) A comparison of currently used methods in computational physical geodesy. This was the primary desire of the contractor.
- (2) A feasibility study on the use of the finite element method for the numerical solution of the fundamental problem of physical geodesy. This was the primary desire of the author.

The fundamental problem of physical geodesy is the simultaneous determination of the earth's figure and potential from geometric and gravimetric measurements. The numerical solution requires a finite parameterization of the potential and - in case of a sophisticated approach - also of the earth's figure.

A comparison of various methods for the detailed representation of the earth's gravity field has recently been given by Tscherning (1979). His confirmed impression is that there is a number of competing methods performing about equally well as far as the quality of results is concerned.

In chapter 2 of the present report various methods currently used to approach the problem of the determination of the earth's figure and potential were examined from the viewpoint of computational efficiency. Methods like collocation, surface layer, buried masspoints, Bjerhammar's method, lead to a fully occupied linear system of equations to be solved. The effort to solve such a system is proportional to N^3 , where N is the number of equations. Breakdown due to OSU-CPU times exceeding 100 hours occurs at about $N = 10,000$ (this corresponds to a surface layer solution with $2^\circ \times 2^\circ$ blocks near the equator).

If the pattern of data and weights shows rotational symmetry, great savings in computation time can be obtained by using techniques

based on discrete Fourier transform of block circulant matrices. This has been shown by Colombo (1980). The author feels that Colombo's approach is currently the best one if an essentially nonredundant set of surface data is employed. Such a set is for example given by $1^\circ \times 1^\circ$ block averages of gravity anomalies. Although the quality of such block averages varies greatly between areas, the assumption of equal weights will not cause too much harm to the estimated parameters, because there is no problem of adjusting redundant data. The system must take what it gets and has no choice to balance poor anomalies against better observations. Of course, the accuracy estimates obtained from such a procedure are very problematic.

Similar things may be said about GEOFAST developed by TASC. The asymptotic speed is even proportional to $N \log N$. The gain in speed is paid for by restricting applications to data distributed regularly on a line or within a rather small plane rectangle. Some possible trouble spots are indicated in chapter 2. One of them is concerned with transporting a covariance from the sphere to the plane. Harmonicity gets lost thereby. It would also be interesting to have some idea on the proportionality factor in front of the $N \log N$ term estimating the CPU time.

Chapters 3 to 6 document a feasibility study on the use of the finite element method in physical geodesy. This method leads to a sparse set of equations whose solution requires an effort proportional to $N^{\frac{3}{2}}$, where N has the same meaning as above. Unfortunately the constant of proportionality is large. The break even point between the surface layer and finite elements in a global solution is estimated to be around $2^\circ \times 2^\circ$ blocks. For smaller blocks finite elements are faster; for larger ones the surface layer is faster. The effort for a global solution based on $1^\circ \times 1^\circ$ gravity anomaly data is estimated at 700 OSU CPU hours. A special technique exploiting the remote zone effect could reduce this to about 250 hours (the surface layer method would require 15,000 hours). An effort of 250 OSU CPU hours is considered too large. Several reruns

would be necessary before a satisfactory choice of weights is found. Although there exist computers, as for example the ILLIAC IV, on which the CPU time could be cut by a factor of about 64, the problem appears too large for one individual researcher or a small research group.

Fortunately, the remote zone effect allows to compute local solutions. The report will give an estimate for the calculation of a detailed potential in an equatorial strip (6.5 hours) and in a rectangular area of size $32^{\circ} \times 64^{\circ}$ (covering e.g. the contiguous US). In the latter case $30' \times 30'$ data were assumed. CPU time was estimated at 20 OSU hours. By a sophisticated use of the remote zone effect this can probably be lowered to 6 hours. This compares favorably with a surface layer solution requiring about 70 hours.

The finite element method does not rely on any regular pattern of observation and weights. (Regularity could be exploited in the same way as with the other methods. The additional saving in CPU time would, however, not be dramatic.) In areas where the field shows much detail, smaller elements may be chosen. Redundant data, as for example gravity anomalies plus geoid heights, pose no problem. The method offers also disadvantages. Harmonicity of the calculated field is only approximate. The programming effort for an efficient computer implementation is considerable.

The finite element method loses much of its efficiency if the data are not local. Local data are composed of measurements taken in a way that one measurement involves only a single point or a small vicinity of a point. A vicinity of a point is considered small if it contains only a small number of the finite elements in its interior. A more precise definition of the locality of a measurement would be that its contribution to the normal equations must not destroy the sparsity pattern resulting from a field representation by means of finite elements. Data obtained by integrating over an unknown orbit are not local. Neither are misclosures of large inertial navigation loops. There

are, however, ways to incorporate orbits at higher altitudes in an efficient way.

The finite element method lends itself to Helmert blocking, or its modern variant, nested dissection. Calculations for subregions (nations, continents) could be delegated. Junction equations could be combined at a higher level, very much in the same way as this is done in continental network adjustment.

Chapter 6 documents a number of test calculations. They were carried out with the following goals.

- (1) To see whether the Ritz-, or the Trefftz-, or the old fashioned least squares principle should be used. (The latter is recommended for the specific needs of physical geodesy).
- (2) To see whether the use of cubic polynomials is sufficient, or whether quintic or even higher degree polynomials are needed. (Cubics are sufficient, quintics are already hopeless from the CPU time point of view).
- (3) To see whether a certain type of element partition making a best possible use of the "attenuation with altitude effect" can be employed. (The outcome was satisfactory).
- (4) To find an appropriate representation of the field in the remote outer space of the earth. (Specially designed elements of infinite size and appropriately chosen shape functions performed well in this respect).
- (5) To get an idea how the observational weights should be balanced against the weights applied to the equations enforcing the approximate fulfillment of Laplace's equation. (Reasonable weights were found by experiment. More insight would be desirable).

(6) To see whether a combination solution of surface gravity values and satellite derived harmonics is possible. (The answer is yes, but additional tests are necessary to identify procedures preventing a substantial increase in CPU time).

The experiments were carried out in 2 dimensions in order to save CPU time. Small scale 3-dimensional calculations are desirable, but there was no time yet to perform them.

In the authors opinion the finite element method has a place in physical geodesy. It is very likely that a proposal by Junkins (1979) will be accepted, suggesting to use a finite element representation of a completely known potential for the purpose of rapid recalculation in real time application and also otherwise. The authors feeling is that finite elements are also useful to parameterise an unknown potential during an estimation procedure. However, the method must be cultivated somewhat more before a large scale effort is attempted. At the end of the research period covered by this report, the author began to look into a hybrid method which combines finite elements with a surface layer of multipoles. Some preliminary statements on this envisioned method are given in chapter 7. Another feature which makes finite elements attractive is the possibility to attack in a head-on way the free boundary value problems of physical geodesy. Some ideas how this could be accomplished are found in chapter 5.

2. Review of various methods

Consider an earth-centered and earth-fixed coordinate system. Choose a convenient reference potential, e.g. that one of an equipotential ellipsoid, or one obtained from a truncated spherical harmonics expansion. The normal potential is assumed to completely absorb the rotational effect. Hence the disturbing potential V is purely gravitational. In outer space it satisfies Laplace's equation

$$\Delta V = 0 \quad (2.1)$$

Laplace's equation represents a local law. In order to evaluate the second order differential operator at a certain location x , we need only information on V in a local neighborhood of x .

The purpose of physical geodesy is the determination of the earth's surface and potential from geodetic measurements. There is hardly a need to point out that the measurements are indirect, and that it is therefore necessary to represent surface and potential in terms of a number of unknown parameters. Because computers perform only finitely many operations in a finite amount of time, the number of parameters must be finite.

The choice of an appropriate set of parameters is highly nontrivial. Factors to be taken into account are:

- (*) Type of application, e.g., local improvements or global corrections to the field.
- (*) Type and distribution of measurements, in particular, homogeneous or heterogeneous sets of data.
- (*) Mathematical simplicity and elegance. Simple setups are easier to program. Debugging the programs is less time consuming.

(*) Computer time during production runs.

Let us restrict attention to the potential, forgetting temporarily the parameters describing the unknown reference surface. Let us review and discuss representations of the potential that have been used by geodesists. Our main emphasis will be on the last item listed above, i.e. computational efficiency during production runs.

2.1. Spherical harmonics.

They satisfy Laplace's equation automatically. The trace functions with respect to the unit sphere are the surface spherical harmonics. They are the eigenfunctions of any rotation invariant operator on the sphere. Confer Mueller (1966), Meissl (1971a), Robertson (1978), Freedden (1979) for extensive discussions. Spherical harmonics provide great theoretical insight. They lead to the concept of the power spectrum of the potential. If a problem can be formulated in a way that rotation symmetry is preserved, then spherical harmonics are also of great computational advantage. In this context recent papers by Freedden (1978), (1979) are pointed out where methods for numerical integration of functions defined on the sphere are specified. The formulas are related to the familiar Gaussian quadrature formulas for intervals, i.e. they rely on a weighted average of function values at a set of discrete points. It is likely that the formulas can be extended to functions discretized in terms of block averages.

Spherical harmonics are widely used in satellite geodesy. At satellite altitudes of 1000 km and above, the potential is sufficiently attenuated to be properly represented by a spherical harmonics expansion of moderately large degree ($N = 20-30$, or so).

Let us now assume that we are dealing with a problem involving a heterogeneous set of data, suggesting a least squares setup by variation of parameters. Then one particular property of spherical harmonics counteracts computational efficiency. This property of spherical harmonics is that they are nonzero almost everywhere. No function different from the zero function can satisfy Laplace's equation and, at the same time, vanish in a part of the domain having nonzero measure. Thus spherical harmonics fail to have a local support. If the disturbing potential is represented in terms of spherical harmonics as

$$V(x) = \sum_{i=0}^N \sum_{j=-i}^{+i} c_{ij} H_{ij}(x) \quad (2.2)$$

and if a local measurement leads to a linear functional $L(V)$ involving only points in a small vicinity of x , then we nevertheless get a representation

$$L(V) = \sum_{i=0}^N \sum_{j=-i}^{+i} c_{ij} L(H_{ij}) \quad (2.3)$$

where most, if not all, of the coefficients c_{ij} are nonzero. Due to the failure of the H_{ij} to have local support, any local measurement will introduce an observation equation into the adjustment which has many nonzero coefficients. The normal equations will be practically full. Solving a symmetric full system of m equations (without complete inversion) requires about

$$\frac{m^3}{6} \quad (2.4)$$

elementary steps, one step comprising one multiplication followed by one addition. Confer equation (A.6) of Appendix A. Assuming that a computer can perform about 500,000 such steps in a second of time, we arrive at the following table, listing CPU times in dependence of various choices for m .

m	CPU time
100	0.3 seconds
1000	5.6 minutes
10000	93.0 hours
100000	10.0 years

Table 2.1. CPU times for
solution of full $m \times m$ systems

$m = 10000$ corresponds to about $N = 100$ in the above expansion for $V(x)$. The actual time required to solve very big systems will be larger than the CPU time due to data transfer between central and peripheral memory.

It may be a coincidence that present day computers limit the spherical harmonics expansion to about $N = 100$. It also appears that the physical significance of spherical harmonics coefficients of degree higher than 100 is questionable. A local anomaly of the field will have a spherical harmonics expansion with c_{ij} tapering off to zero more reluctantly the more pronounced the anomaly is, i.e., the less smooth it is. Hence a local anomaly is decomposed into components being nonlocal. This may be desirable in such fields as optics or acoustics. In physical geodesy it is undesirable.

2.2. Surface layer representation and related methods.

Under sufficiently general conditions the outer potential can be represented by a surface layer. A surface layer can be specified in terms of finitely many parameters in various ways, such as for example by constant values of density in subregions composing the entire surface of the earth. Confer Koch and Witte (1971), Morrison (1980). Any of the surface elements generates a potential, i.e., a solution of Laplace's

equation in outer space. The total potential is obtained by superposition. In this respect, the presently discussed method does not deviate from spherical harmonics. Also the property of a computationally rather undesirable nonlocal support carries over. Hence the above table continues to give an indication of the computational effort involved, if m is taken as the number of surface elements. Indeed, global solutions exceeding $m = 2000$, which corresponds to about 5 by 5 degree elements near the equator, have not been reported in the literature. (Confer, however, subsection 2.7 below dealing with shortcuts resulting from symmetrical configurations).

The physical significance of the surface layer is not immediate. However they definitely offer the advantage of modelling local effects. In areas, where the field is very detailed, or, where a more detailed knowledge of the field is desired, smaller sized elements may be chosen.

Similar statements can be made about buried masspoints (confer e.g. Needham (1970) or Hardy and Goepfert (1975)) and about Bjerhammar's method. The latter represents the potential by its boundary values on a sphere entirely contained in the interior of the earth. A finite parameterization is achieved e.g. by partitioning the surface of the sphere into small elements, and by assuming constant boundary values in these elements. Confer Bjerhammar (1978), Sjöberg (1978).

The normal equation system will be full, i.e. the above table applies. Local effects may be conveniently modelled by varying the element size.

2.3. Least squares collocation.

2.3.1. Krarup's proposal.

Although least squares collocation shares some features with the methods mentioned under 2.2, there are some important deviations. Again, the

potential is represented as a linear superposition of special solutions to Laplace's equation:

$$V(x) = \sum_{i=1}^n c_i L_i K(x, y_i) \quad (2.5)$$

The number of terms, however, now equals n , the number of measurements. The functions

$$F_i(x) = L_i K(x, y_i) \quad (2.6)$$

are derived from a symmetric and positive definite kernel $K(x, y)$ which satisfies Laplace's equation with respect to x . (Due to symmetry, it also satisfies Laplace's equation with respect to y). If the i -th measurement l_i refers to a functional

$$L_i = L_i(V(x)) ; \quad x = x_i \quad (2.7)$$

then $F_i(x)$ as given by equation (2.6) is taken as the i -th basis function.

The function $K(x, y)$ is viewed as a reproducing kernel. Thus it defines a norm $\|V\|$. The c_i are chosen such that

$$L_i(V(x)) = l_i \quad (2.8)$$

and that

$$\|V(x)\| = \text{Min.} \quad (2.9)$$

Confer Krarup (1969). The textbook by Moritz (1980) may be consulted for a detailed documentation, discussion, presentation of extensions, and for its bibliography.

It is seen that least squares collocation uses as many basis functions as there are observations. As compared to 2.1 and 2.2 this number n will frequently be larger than m , the number of unknowns in the other methods. Hence we can expect a very good fit. However, a price has

to be paid for this. The size of the linear system to be solved is n by n . The system is full for the very same reasons as given earlier: the $F_j(x)$'s satisfy Laplace's equation. Hence the computational effort is

$$\frac{n^3}{6} \quad (2.10)$$

The mathematical elegance of least squares collocation is undisputed. The method easily takes any type of heterogeneous data. The choice of a suitable covariance function is not immediate and requires insight. Confer Tscherning and Rapp (1974).

A computer implementation of the collocation method is described in Tscherning (1978).

2.3.2. Least squares collocation using unknown parameters.

A somewhat unsatisfactory aspect of pure least squares collocation is the following one. In areas of insufficient data the predicted function tends to approach the zero function. Since the problem of physical geodesy results from a linearization procedure based on the use of a reference surface, the consequence is that in areas of insufficient data the reference surface is predicted. The reference surface is, however, mostly chosen according to computational convenience rather than according to its approximation of the physical truth.

This unsatisfactory aspect can be counteracted by using a setup including unknown parameters. Confer Moritz (1980), section 16. This setup is closely related to the concept of generalized splines which will be discussed farther below.

2.4. Finite elements

Finite elements have been very successful in other disciplines. There exists an abundance of literature. As textbooks we mention Zienkiewicz (1971), Schwarz (1980), Ciarlet (1978), Strang-Fix (1973). Finite elements have also been used by some geodesists. Cf. Szameitat (1979), Werner (1979). Bosman-Eckhart-Kubik were early geodetic users of finite element concepts. They applied piecewise polynomials to surface approximation problems. The use of finite elements in Physical Geodesy (in the narrow sense) was up to now restricted to the representation of a known potential for rapid recalculation. Confer Junkins (1977), (1979), Engels (1979). We intend to use the method also during the determination of the potential together with the earth's surface. Our intended use will be described in detail in the subsequent chapters. In this section we shall be very brief.

The domain of interest is subdivided into finitely many subregions, called elements, of preferably simple shape. If the region is unbounded, some of the elements must be of infinite size. We shall mostly work with box-type elements partitioning the r, φ, λ parameter space resulting from a choice of polar coordinates. At the boundary of any element a number of nodes is located. Any node is shared by two or more elements. In our case nodes will be mostly at the corners of the boxes; but occasionally some are also encountered elsewhere on the faces.

To any node a number of parameters is associated. Usually they include the value of the potential there and of some of its derivatives. An interpolation formula is prescribed which allows to calculate the potential and its derivatives at any point in the interior or on the boundary of an element from the parameter values of all nodes associated with this element. The interpolating function is analytic and of simple

shape in the interior of the elements. Across element boundaries continuity of the function is usually required. Depending on the type of application, continuity of some derivatives is also needed. We propose the use of tricubic polynomials in the elements such that the resulting potential is globally C^1 continuous (it is continuous together with its first order derivatives).

If an observation of the potential refers to a location within an element, the resulting observation equation will involve only parameters of nodes associated with this element. As a consequence, the system of observation equations will be sparse and so will be the normal equations formed from them. This is a great computational advantage. The normal equations resulting from the observation equations are not yet sufficient. A potential represented by finite elements is not automatically harmonic. Harmonicity must be enforced by another set of normal equations which must be added to the earlier ones. We shall call this the contribution of the field to the normal equations. It is obtained by minimizing the integral over the square of the Laplacean of the field. Harmonicity is not fully ensured this way, but only approximately. The sparse structure of the normal equations is not impaired by the field contribution. Sparseness is the great benefit of the finite element method. If N is the total number of parameters in our application of the method, the computational effort will be

$$\text{const } N^{\frac{3}{2}} \quad (2.11)$$

As we shall see, the total number of parameters is appreciably larger than the number of blocks in a comparable surface layer solution. However it is important to stress that a synchronized refinement of the partition in the two methods will keep the ratio of the number of blocks to the number of parameters approximately constant. Hence our method is $\text{const}_F * N^{\frac{3}{2}}$ as compared to $\text{const}_S * N^3$ in the surface layer solution. On the

other hand $\text{const}_p > \text{const}_s$. As a consequence, for a small number of blocks the surface layer solution will be more economical. For large N , the finite element solution is better. The break even point is estimated to be around $2^\circ \times 2^\circ$ blocks.

Remark: A computational effort of $O(N^{\frac{3}{2}})$ steps results if the "nested dissection method" due to George (1973), (1977) is applied to a sparse system of N equations resulting from decomposing a two dimensional region of size $O(\sqrt{N}) \times O(\sqrt{N})$ into N elements. The system represents the equilibrium equation for a 2-dimensional elastic problem defined over this region. Although our region is 3-dimensional, the estimate of $O(N^{\frac{3}{2}})$ steps remains valid due to the opportunity to use increasingly larger elements as the altitude increases.

Remark: it must be emphasized that the efficiency of the finite element method relies heavily on the locality of the measurement functionals. Any measurement must involve only one point or a set of points confined to a small region. Measurements involving points along an unknown trajectory, such as for example misclosures of large inertial navigation loops, are excluded. They would destroy the sparsity pattern and degrade the asymptotic computational efficiency to that of least squares collocation. Also unknown orbits of satellites pose difficulties. If the orbits are high enough, one may, however compromise by fusing finite elements near the earths surface with e.g. spherical harmonics at higher altitudes. The elements are chosen small near the earths surface. They get larger and larger with altitude in agreement with the potentials attenuation. At satellite altitude there is only one element, or a small number of them.

2.5. Spline functions.

Theory and application of spline functions are very diversified. There is an overlap with least squares collocation and a border line with finite elements. Spline functions were invented by I. J. Schoenberg and first described in his famous paper Schoenberg (1946). Since then they have evolved into a very popular tool of applied mathematicians as well as into an object of interest to theoreticians, who implanted them into Hilbert spaces. Textbooks have been published, as for example Ahlberg et. al. (1967), Boehmer (1974).

The practically minded person associates with splines a special subset of them, namely polynomial splines. There is a widespread preference for cubic splines. Polynomial splines perform well in interpolation problems due to their simplicity, computational efficiency, smoothness and locality. As already pointed out by Schoenberg (1946), one can construct basis functions having a local support.

The use of spline functions for problems of physical geodesy was suggested by Davis and Kontis (1970). Meissl (1971b) proposed their use for the representation of pointwise known functions during the evaluation of the explicit integral formulas of physical geodesy, i.e. the formulas by Stokes, Vening Meinesz and their refinements due to Molodensky. This proposal was worked out by Suenkel (1977) and Noë (1980).

From the point of view of Hilbert space theory spline functions are optimal interpolators (or approximators) of functionals. Optimality relies on two complementary criteria, namely the minimum norm property and the best approximation property. Both criteria are based on the choice of a seminorm. This choice is up to the user. In contrast to general least squares collocation, certain natural curvature-seminorms strongly suggest themselves as candidates.

Theory and geodetic use of splines are discussed in detail in Moritz (1978), Lelgemann (1980) and in a forthcoming paper by Freeden (1981). Collocation with the use of parameters can be inbedded

into the Hilbert space theory of splines as outlined in Boehmer (1974), chapter 4.

We shall briefly stress the point of view of computational efficiency. If one uses generalized splines, as proposed by Leigemann (1980), Freedman (1981), one deals with functions lacking a local support. Hence the normal equations are full. This limits the size of the systems to a few thousand unknowns. On the other hand cubic splines can be generalized to 2 and 3 dimensions by means of tensor products. Here basis functions with local support are available. Thus the spline method competes with the finite element method in computational efficiency. Therefore we shall discuss this particular point in some detail.

Imagine, for simplicity, a rectangular region in 3-dimensional space subdivided into box-type elements of equal size and shape. The set of nodes shall be identical with the set of corners of the element. Our intended use of the finite element method relies on interpolating functions called Hermite tri-cubics. In any of the various boxes, the function to be interpolated is represented by a polynomial which is a cubic in any one of the 3 variables, provided that the other two variables are fixed. (Considered as a polynomial in 3 variables the interpolating function is of degree 9). We associate with any node 8 parameters representing

$$\frac{\partial^i}{\partial r^i} \frac{\partial^j}{\partial y^j} \frac{\partial^k}{\partial \lambda^k} V(r,y,\lambda) ; \quad 0 \leq i,j,k \leq 1 \quad (2.12)$$

at this node. These are the derivatives of the potential of "bidegree" less than or equal to 1. By letting all parameters having the value zero except for one, we obtain as interpolating functions a basis function associated with this particular node and this particular parameter. This basis function is called shape function. This will be discussed in detail in section 3.2. Here we only emphasize that we have 8 basis functions per node. Any basis function is C^1 continuous and has a local

support limited to 8 adjacent elements

If we alternatively represent our field by means of basis functions built up of tri-cubic splines, we obtain only 1 basis function for each node. It is the tensor product $B(x,y,z) = B(x)B(y)B(z)$ of one-dimensional B-splines $B(x)$ specified in Schoenberg (1946), p. 71, and called $M_4(x)$ there. This basis function $B(x,y,z)$ is even C^2 continuous. This makes splines attractive as compared to Hermite tri-cubics. However, it turns out that the support of any basis function covers 64 adjacent elements

Let us summarize and conclude the comparison of finite elements and splines by the following 3 statements.

(1) Tricubic splines have basis functions involving less parameters than Hermite tricubics. This may be viewed beneficial. On the other hand it also means that we have less flexibility unless we decrease the size of the elements.

(2) Splines are smoother. This makes them more useful in interpolation problems. However, in problems of representing a field governed by a differential equation, we have an additional enforcer of smoothness. This was called the field contribution to the normals in section 2.4. For this reason splines are preferred in pure interpolation problems, whereas Hermite polynomials are preferred in the finite element solution of field equations. Confer the discussion in Strang-Fix (1973), p. 61.

(3) Due to the larger support of splines, the linear system will be less sparse. In an oversimplified way, we may talk of a larger bandwidth as compared to a system resulting from the use of Hermite tricubics. It appears that, whatever may be gained by a decreased number of parameters and by greater smoothness in case of splines, is paid for by a larger bandwidth slowing down the elimination procedure.

2.6 Approximate explicit Green's functions

Frequently the oldest methods are the best. Hardly ever they are the worst methods computationally. Computers were unavailable at earlier times. Take Stokes' formula. It yields the geoidal undulation at one point in terms of gravity anomalies all over the globe. It is hardly necessary to point out the approximations underlying Stokes formula as well as the corrections which partly make good for them. The usefulness of Stokes formula as well as of Vening Meinesz formula and their refinements is undoubted. Applications are, however, restricted to areas of moderately varying topography.

Stokes' and Vening Meinesz' kernel are explicitly known Green's functions of boundary value problems for the sphere. We are in a similar situation as, when dealing with a large system of linear equations, an a-priori known inverse of the coefficient matrix is available. If the boundary value problem is discretized in agreement with a discretization of the gravity anomalies in terms of N block averages, we obtain indeed such a system of N linear equations in N unknowns. If the inverse is known, calculation of the solution requires N steps for one particular unknown and N^2 steps for all of them. This is not impressive in itself because we know that in case of a sparse system of the type mentioned in subsection 2.4 we can do better, namely solve for all unknowns in $O(N^{\frac{3}{2}})$ steps. In case of evaluating the discretized Stokes formula, an important additional bonus is available, namely the remote zone effect. It implies that of the N steps necessary to evaluate one specific unknown, many can be lumped into comparatively few new steps, and many may even be omitted altogether. The number of new steps to be carried out is a fraction αN of N , where α is viewed as a fixed constant. The constancy of α is based on the following argument. Suppose that a certain block design is used for the approximate evaluation of Stokes integral. Near the point of evaluation we use averages of gravity

anomalies over the smallest blocks available, say $1^{\circ} \times 1^{\circ}$ blocks. This corresponds to $N = 360 \times 180 = 64800$. At a moderate distance we may lump blocks to $2^{\circ} \times 2^{\circ}$ and so on. Very distant blocks may be omitted, in particular in case of Vening Meinesz' formula. If we quadruple N , proceeding from $1^{\circ} \times 1^{\circ}$ blocks to $30' \times 30'$ blocks, any of the lumped blocks in the above design is split into 4 new blocks. Hence the number of steps also quadruples.

It appears that the effort needed to calculate the geoid at one point (block center) is $\propto N$, and $\propto N^2$ for N points. Hence the method is still $O(N^2)$, but the constant hidden under the " O "-symbol is very small. Finite elements are $O(N^{\frac{3}{2}})$, and consequently asymptotically better. However, the constant hidden in $O(N^{\frac{3}{2}})$ is large. The break even point is not exactly known now.

If geoid or deflection of the vertical are needed only at one point or at a very small number of points, the explicit inverse method is the best, namely $O(N)$ with a very small hidden constant. However it must be stressed that the explicit inverse method relies on a special type of homogeneously distributed and nonredundant data. There is no way to vary the weights individually for the blocks. Additional data are difficult to incorporate in a theoretically satisfactory way.

2.7. Exploiting rotational or translational symmetries.

2.7.1. Invariance of normal equations under a group of transformations.

We are not referring to a method that stands for its own as those described thus far. We are dealing with a technique that can be applied in conjunction with any of the methods described in subsections 2.2 to 2.6, provided that the distribution of data satisfies certain requirements which are rather stringent.

If the pattern of measurement-locations and -weights is invariant with respect to the group of rotations around an axis, or with respect to the group of translations in 1, 2, or 3 dimensions, then this symmetry is reflected by the system of normal equations to be solved. Of course, proper care must be taken, that the parameterization of the potential conforms with the symmetry. The normal equation system is invariant with respect to one or more transformations generating the group, provided that the unknowns are properly renumbered. Let

$$Gx = r \quad (2.13)$$

denote the original normal equations. Let

$$x = Hy \quad (2.14)$$

be the transformation taking into account the translation or rotation followed by a renumbering. The transformation will be orthogonal, i.e.,

$$H^T H = I \quad (2.15)$$

The new normal equations are

$$H^T G H y = H^T r \quad (2.16)$$

They are identical to the old ones. Hence

$$H^T G H = G \quad \text{or} \quad G H = H G \quad (2.17)$$

Two matrices which commute share a common system of eigenvectors. It follows that an invariant subspace of H is also an invariant subspace of G . Invariant subspaces of H are usually easy to identify. H reflects only the symmetries of the problem and is independent of other structural properties. The knowledge of invariant subspaces allows the decomposition of the normals (2.13) into several independent systems of smaller dimension.

2.7.2. Outline for the case of rotational symmetry around an axis

Assume that the system is invariant with respect to a rotation around one axis by an angle of

$$\beta = \frac{2\pi}{K}$$

This case arises, if 5 degree by 5 degree mean gravity anomalies are taken as measurements and if the field is parameterized by a surface layer with constant density in $N = 2592$ blocks of size $5^\circ \times 5^\circ$. We then have $k = 360/5 = 72$. Imagine the parameters (block densities) grouped according to longitude. The groups are numbered according to increasing longitude. For a certain fixed longitude, we imagine a numbering according to decreasing latitude. Note that a rotation by β carries all blocks of a certain longitude λ over into blocks of longitude $\lambda + \beta$. Hence a cyclic renumbering of the groups of blocks is necessary in order to ensure invariance of the normal equations under the transformation H . H will be of the form

$$H = \begin{bmatrix} & & & & I \\ & & & & \\ I & & & & \\ & I & & & \\ & & I & & \\ & & & \ddots & \\ & & & & I \end{bmatrix} \quad (2.18)$$

The size of the diagonal blocks I is implied by the number of blocks having the same longitude. Invariant subspaces of H , which is a permutation matrix, are immediately specified. They are given by the k block-columns of the following unitary matrix

$$Z = \frac{1}{\sqrt{k}} \begin{bmatrix} Z_{00} & Z_{01} & \dots & Z_{1,K-1} \\ Z_{10} & & & \vdots \\ \vdots & & & \vdots \\ Z_{K-1,0} & \dots & \dots & Z_{K-1,K-1} \end{bmatrix} \quad (2.19)$$

with

$$Z_{rs} = e^{-i \cdot rs \frac{2\pi}{k}} I \quad ; \quad r, s = 0, 1, \dots, k-1 \quad (2.20)$$

$$I \dots \text{unit matrix} \quad ; \quad i = \sqrt{-1}$$

A parameter transformation

$$x = Z z \quad (2.21)$$

will decompose the normal equation system into k independent systems of size N/k . Their solution will require an effort proportional to

$$\frac{k}{6} \left(\frac{N}{k} \right)^3 \quad (2.22)$$

Since $k = O(\sqrt{N})$, this effort is $O(N^2)$. Of course, also the computations required to transform the system must be taken into account. However here one may employ fast Fourier transform techniques.

Transformations like that one outlined above have been used before in other disciplines. They have been used in geometrical geodesy by Meissl (1969) in order to analyse the strength of regular triangulation chains. Their first use in physical geodesy is due to Colombo (1980).

2.7.3. The effort of the TASC group.

It should be mentioned that translational symmetry requires observations covering an infinite line or an infinite plane. If the domain is restricted to a finite rectangle, boundary effects destroy the symmetry. Nevertheless, one is able to salvage most of the saving encountered in the undisturbed case. The methods are more involved. They have been widely used in picture processing. Recently an effort has been made by

TASC (= The Analytic Science Corporation) to analyse geophysical data distributed regularly on an arbitrarily long line or within a small rectangle. Confer Heller, Tait and Thomas (1977), Tait (1979)

The computational effort in both cases is proportional to $N \log N$ where N is the size of the linear system to be solved. Hence N is equal to the number of unknowns in a parameter model or equal to the number of measurements in a collocation model. The approach of the TASC group is interesting. It rests on two techniques, namely (1) the Fast-Fourier-Transform (due to Cooley and Tukey (1965)) and (2) on the block decomposition of block circulant matrices as outlined above. In addition to these two ingredients, the authors employ a number of tricky maneuvers in order to tackle the above mentioned undesirable boundary effect. By using a transition from an $N \times N$ Toeplitz matrix to a $2N \times 2N$ block circulant matrix, and by employing a data window, they finally arrive at a linear system in transformed space where a few diagonals near the main diagonal are strongly dominant. After neglecting the element outside this band, and after adding a small multiple of the unit matrix to the coefficient matrix, the banded system is solved by Cholesky.

We are unable to present the details in this report and refer the reader to the quoted original articles. We only make the following 4 remarks.

(1) The solution is approximate, even if the method is applied to regular data on a straight line segment or in a plane rectangle. The errors come from two sources, namely (a) the neglect of elements outside the band and (b) the addition of the small multiple of the unit matrix. The procedure (b) results in deemphasizing the weights of observations near the boundary of the region. The authors give error estimates which are favorable. However, it is not clear whether such favorable error estimates are available for other situations than those

considered by the authors. They test their method in predicting gravity anomalies from geoid heights by a collocation procedure. Thus they deduce a high frequent output from a low frequent input. (Cf. Meissl (1971) for a discussion of the frequency content of various quantities related to the earths disturbing potential). It should also be tested whether deducing a low frequent output from a high frequent input can be done with errors of the same small magnitude.

(2) In the case of a two dimensional area, the method works strictly speaking only for a plane region. Mapping a part of the sphere (spheroid) onto the plane causes distance distortions, as the authors point out. However they do not point out that these distance distortions interact with the covariance kernel, causing it to fail to fulfill Laplace's equation any longer. Harmonicity of the covariance function is an inherent assumption in collocation. The authors map the systems of meridians and parallels onto a rectangular grid in the plane. Such a mapping has appreciable distance distortions. It is well known that there are mappings which perform better in this respect. Perhaps one of them should be used.

(3) Intuitive insight into the method can be gained by the following consideration. Consider the familiar collocation problem for two random vectors:

$$y = C_{yx} C_{xx}^{-1} x \quad (2.23)$$

We assume that x and y are related to discrete equidistant points $i = 0, 1, \dots, n$ on the line. We assume homogeneous processes, hence C_{xx} is a symmetric Toeplitz matrix:

$$C_{xx} = \{C_{ij}\} = \{C_{|i-j|}\} \quad (2.24)$$

We assume that C_{xx} is of the form

$$C_{xx} = C + D \quad (2.25)$$

where C is positive definite Toeplitz and D is diagonal and positive definite. D is due to measurement noise. The main problem is the calculation of

$$z = (C + D)^{-1} x \quad (2.26)$$

i.e. the solution of

$$(C + D)z = x \quad (2.27)$$

As TASC proposes, we extend C to a $2n \times 2n$ Toeplitz circulant matrix \tilde{C} where

$$\tilde{C}_k = \begin{cases} C_k & 0 \leq k \leq n-1 \\ 0 & k = n \\ C_{|2n-k|} & n+1 \leq k \leq 2n-1 \end{cases} \quad (2.28)$$

D will also be diagonally extended to \tilde{D} as we show in a moment. We extend x by zeroes

$$\tilde{x} = \begin{pmatrix} x \\ 0 \end{pmatrix} \quad (2.29)$$

and thus obtain the system

$$(\tilde{C} + \tilde{D})\tilde{z} = \tilde{x} \quad (2.30)$$

This system is not equivalent to the earlier one (2.27), in the sense that the restriction of \tilde{z} to the first n components is the solution of (2.27). The reason for the failure of (2.30) is the following one. The system (2.30) predicts y not only from the measured x_i ; $i = 0, 1, \dots, n-1$, but also from the artificially assumed values $x_i = 0$; $i = n, \dots, 2n-1$. Hence the prediction is theoretically wrong. However there is

still the matrix \tilde{D} which can help us to make the prediction nearly correct. This can be accomplished by assuming very large elements of \tilde{D} at the positions $i = n, \dots, 2n-1$. This amounts to the superposition of very heavy noise on the artificially assumed measurements $x_i = 0, i = n, \dots, 2n-1$. Hence (2.30) will lead to a nearly correct prediction.

We now apply the discrete Fourier transform toward (2.30) Introducing the $2n \times 2n$ Fourier matrix

$$\tilde{F} = \{ \tilde{F}_{jk} \} ; \quad \tilde{F}_{jk} = \frac{1}{\sqrt{2\pi}} e^{-\frac{2\pi i j k}{2n}} ; \quad i = \sqrt{-1} \quad (2.31)$$

the system (2.30) goes over into

$$(\tilde{F} + \tilde{\Delta}) \tilde{y} = \tilde{y} \quad (2.32)$$

with

$$\begin{aligned} \tilde{F} &= \tilde{F}^H \tilde{C} \tilde{F} \\ \tilde{\Delta} &= \tilde{F}^H \Delta \tilde{F} \\ \tilde{y} &= \tilde{F}^H \tilde{x} \\ \tilde{z} &= \tilde{F}^H \tilde{z} \end{aligned} \quad (2.33)$$

F^H is the Hermitian adjoint of F . Hence F^H is also equal to F^{-1} .

\tilde{F} is now diagonal. This is the benefit from applying Fourier transformation to a Toeplitz circulant matrix. Cf. also the discussion around equation (2.20)

$\tilde{\Delta}$ is a Toeplitz circulant. Hence it appears doubtful presently what we have gained. However one can verify that the subdiagonal bands of $\tilde{\Delta}$ decrease as one goes away from the main diagonal. This is perhaps intuitively clear if one remembers that the Fourier coefficients of the

step function

$$S(t) = \begin{cases} 0 & 0 \leq t \leq \pi \\ 1 & \pi \leq t \leq 2\pi \end{cases} \quad (2.34)$$

are given by

$$\frac{1}{2\pi n i} (1 - (-1)^n) \quad (2.35)$$

The tapering effect of the subdiagonal bands can be made more pronounced if the transition of small elements of D in $[0, n-1]$ to large ones in $[n, 2n-1]$ is smoothed out somewhat. This causes "data deemphasis" of x_i near the interval ends.

The next step is to neglect the off diagonal bands of \tilde{A} up to a small number m ($m = 5$ to 10 , or so). This truncated banded system version of the system (2.32) is now solved by Cholesky in $O(n)$ steps. The rest can be accomplished in $O(n \log n)$ steps if the Fast Fourier technique is employed.

We leave it with this oversimplified picture of the algorithm which could be extended to the 2-dimensional case. It is interesting to note that the above mentioned "data deemphasis effect" is also encountered in the presentation of the real TASC algorithm.

(4) Shortly before finishing this report, an article by Bitmead-Anderson (1980) came to my attention. The authors show how an $n \times n$ Toeplitz system can be solved in $O(n \log^2 n)$ steps by a doubling method. The matrix is subjected to some mild restrictions, however reference is made to other work by Brent et. al. (1980), in which an $O(n \log n \log n)$ algorithm is specified which achieves a solution whenever it exists. It should be noted that no approximations are involved as they are in the work of TASC. Of course, the question arises again how large the constant hidden in the "O" symbol really is.

2.7.4. Rauhala's array algebra.

Another way to utilize symmetries was pointed out by Rauhala. Confer Rauhala (1980) for details and further references. Also Snay (1978) gives an introduction to Rauhala's "array algebra". It is an application of the concept of multilinear mappings between tensor spaces. Let X, Y, Z denote 3 vector spaces of not necessary equal dimension. Consider the space T of tensors:

$$T = X \otimes Y \otimes Z \quad (2.36)$$

An element of this space is represented by a three-dimensional array $t_{i,j,k}$. T can be viewed as the linear span (set of linear combinations) of vectorial tensor products $x \otimes y \otimes z$ with $x \in X, y \in Y, z \in Z$. The tensor generated by $x \otimes y \otimes z$ has elements $t_{i,j,k} = x_i y_j z_k$. Consider three further vector spaces X', Y', Z' of arbitrary dimensions I', J', K' . Form the tensor product $T' = X' \otimes Y' \otimes Z'$. Let A, B, C be linear operators

$$\begin{aligned} x' &= Ax \\ y' &= By \\ z' &= Cz \end{aligned} \quad (2.37)$$

Define a linear map $T \rightarrow T'$ in the following way. Let the image of $x \otimes y \otimes z$ be $Ax \otimes Ay \otimes Az$. Extend the domain from the set of tensor products to all of T by means of linearity. Thus a map $A \otimes B \otimes C$ from $T \rightarrow T'$ is obtained. Suppose, temporarily, that any of the maps A, B, C is invertible. It is easily proved that

$$(A \otimes B \otimes C)^{-1} = A^{-1} \otimes B^{-1} \otimes C^{-1} \quad (2.38)$$

Here the benefit from array algebra becomes transparent: instead of inverting a huge matrix of size $(IJK) \times (IJK)$, one inverts 3 matrices of

sizes $I \times I$, $J \times J$, $K \times K$.

Assume now $I' > I$, $J' > J$, $K' > K$ and consider the least squares problem

$$t' + v = (A \otimes B \otimes C) t \quad (2.39)$$

with $t \in T$, $t' \in T'$, and $v \in T'$ denoting the residuals. The 3 least squares problems

$$\begin{aligned} x' + v_x &= Ax \\ y' + v_y &= By \\ z' + v_z &= Cz \end{aligned} \quad (2.40)$$

are solved by the pseudo-inverses

$$\begin{aligned} x &= A^+ x' \\ y &= B^+ y' \\ z &= C^+ z' \end{aligned} \quad (2.41)$$

If the rank of A equals the number of its columns, then

$$A^+ = (A^T A)^{-1} A^T \quad (2.42)$$

and similarly for B^+ , C^+ .

Rauhala shows that the pseudo-inverse of $A \otimes B \otimes C$ is given by

$$(A \otimes B \otimes C)^+ = A^+ \otimes B^+ \otimes C^+ \quad (2.43)$$

A proof follows easily from geometric reasons based on "range-space" and "null-space" considerations. We do not give a complete proof here, because it requires a number of formal definitions. We merely mention the following facts:

(a) The matrix A maps its domain space one to one onto its range space.

The orthocomplement of the domain space is the null-space.

(b) The pseudoinverse maps the range space inversely back onto the domain space. It maps the orthocomplement of the range space onto zero.

(c) Domain space and range space of $A \otimes B \otimes C$ are the tensor products of domain and range spaces of A , B , C . Hence (2.43) is essentially reduced to (2.38).

Thus the least squares problem (2.39) is solved by

$$t = (A^+ \otimes B^+ \otimes C^+) t' \quad (2.44)$$

Let this outline be enough. We just mention that generalization to tensor products of arbitrarily many factors are immediate. Rauhala will also forgive that I did not use the most generalized inverses he has ever invented.

Applications of array algebra are restricted to gridded problems defined on regions being of the box-type. The grids must be rectangular, however the spacing between gridlines may vary. The question still remains how familiar problems of physical geodesy are transformed into problems of array-algebra. Rauhala states that this cannot be done without some "cheating a la Gordian knot". The cheating may perhaps be comparable to that one encountered during the transformation of a problem formulated for a small spherical rectangle into a translation invariant problem defined over a plane rectangle.

Let us discuss the computational effort for the case of two vector spaces X , Y of equal dimension n . The tensor product $T = X \otimes Y$ has dimension $N = n^2$. The effort to naively solve an $N \times N$ system requires $O(N^3) = O(n^6)$ steps. The effort to invert 2 matrices of $n \times n$ is $O(n^3) = O(N^{\frac{3}{2}})$. It may be shown that a solution of the system utilizing the

decomposition $A \otimes B$ can also be done in $O(N^3)$ steps. This is asymptotically the same effort as if a general sparse system resulting from a 2-dimensional layout is solved by the nested dissection method. Confer the discussion in section 2.4, in particular the first of the 2 remarks given at the end of subsection 2.4.

3. Outline of the finite element approach.

3.1. Hermite-cubic representation of the field

3.1.1. The four basis functions for the unit interval

Let

$$\begin{aligned}\chi(x) &= (x-1)^2(2x+1) \\ \omega(x) &= x(x-1)^2\end{aligned}\quad 0 \leq x \leq 1 \quad (3.1)$$

and define

$$\begin{aligned}\gamma_{00}(x) &= \chi(x) &= 2x^3 - 3x^2 + 1 \\ \gamma_{01}(x) &= \omega(x) &= x^3 - 2x^2 + x \\ \gamma_{10}(x) &= \chi(1-x) &= -2x^3 + 3x^2 \\ \gamma_{11}(x) &= -\omega(1-x) &= x^3 - x^2\end{aligned} \quad (3.2)$$

Graphs of these 4 basis functions are shown in figure 3.1

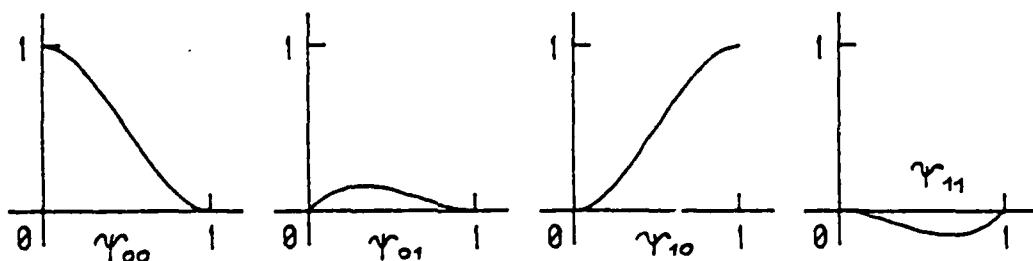


Fig. 3.1 Graphs of basis functions in unit interval

Note that the first subscript in $\gamma_{ij}(x)$ refers to the location, i.e. $x = 0$ or $x = 1$. The second subscript refers to the degree of the derivative which is equal to unity at that node.

The 4 basis functions solve the following interpolation problem:

given

$$\begin{aligned} f(0) &= C_{00} & f(1) &= C_{10} \\ f'(0) &= C_{01} & f'(1) &= C_{11} \end{aligned} \quad (3.4)$$

find a cubic polynomial interpolating these values. The solution is

$$f(x) = \sum_{i,j=0}^1 c_{ij} \gamma_{ij}(x); \quad 0 \leq x \leq 1 \quad (3.5)$$

We also introduce the derivatives

$$\gamma_{kij}(x) = \frac{d^k}{dx^k} \gamma_{ij}(x); \quad k = 0, 1, 2 \quad (3.6)$$

Remark: The coefficients of

$$\gamma_{kij}(x) = \sum_{\ell=0}^3 c_{kij\ell} x^\ell \quad (3.7)$$

are stored in the 4-dimensional array PSC(K,I,J,L) during execution of our computer programs described in chapter 6

3.1.2. Interval of length h

Let

$$\gamma_{ij}(x;h) = h^j \gamma_{ij}\left(\frac{x}{h}\right); \quad 0 \leq x \leq h \quad (3.8)$$

These functions solve the above interpolation problem with data given at $x = 0$ and $x = h$. Again we take derivatives of these basis functions:

$$\gamma_{kij}(x;h) = \frac{d^k}{dx^k} \gamma_{ij}(x;h); \quad k = 0, 1, 2 \quad (3.9)$$

Obviously:

$$\gamma_{kij}(x;h) = h^{j-k} \gamma_{kij}\left(\frac{x}{h}\right) \quad (3.10)$$

3.1.3. Bicubic polynomial in a rectangle with sides a, b

Take

$$\begin{aligned} \gamma_{i_x i_y j_x j_y}(x,y;a,b) &= \gamma_{i_x j_x}(x;a) \cdot \gamma_{i_y j_y}(y;b) \\ 0 \leq x \leq a \\ 0 \leq y \leq b \end{aligned} \quad (3.11)$$

Note: the first two indices refer to the location, the second pair refers to the degree of derivatives!

The functions (3.11) solve the following interpolation problem:

Given

$$\begin{aligned} f(0,0) &= c_{0000} & f(0,b) &= c_{0100} \\ f_y(0,0) &= c_{0001} & f_y(0,b) &= c_{0101} \\ f_x(0,0) &= c_{0010} & \vdots & \\ f_{xy}(0,0) &= c_{0011} & f_{xy}(a,b) &= c_{1111} \end{aligned} \quad (3.12)$$

find a bicubic polynomial

$$f(x,y) = \sum_{r=0}^3 \sum_{s=0}^3 a_{rs} x^r y^s \quad (3.13)$$

interpolating these data. The solution is

$$f(x,y) = \sum_{i_x, i_y, j_x, j_y} c_{i_x, i_y, j_x, j_y} \gamma_{i_x, i_y, j_x, j_y}(x,y;a,b)$$

$$0 \leq x \leq a$$

$$0 \leq y \leq b$$
(3.14)

3.1.4. Tricubic polynomials in a box with sides a, b, c

The extension to 3 dimensions is obvious: In the interpolation problem we prescribe the nodal derivatives shown in table 3.1 at the 8 corners of a box with side-length's a, b, c

j_x, j_y, j_z	000	001	010	011	100	101	110	111
Derivative	f	f_z	f_y	f_{yz}	f_x	f_{xz}	f_{xy}	f_{xyz}

Table 3.1

Numbering of nodal derivatives in 3 dimensions

3.1.5. C^1 continuity across element boundaries.

Consider an $n = 1, 2$ or 3-dimensional region divided into elements. For $n = 1$ the elements are intervals, for $n = 2, 3$ the elements are rectangular boxes. For $n = 2, 3$ we assume that a corner of an element does not touch the interior of a face (boundary segment, boundary rectangle) of an adjacent element. Otherwise, the elements need not be of the same size. Take as an example the partition of a region in R^2 shown in figure 3.2.

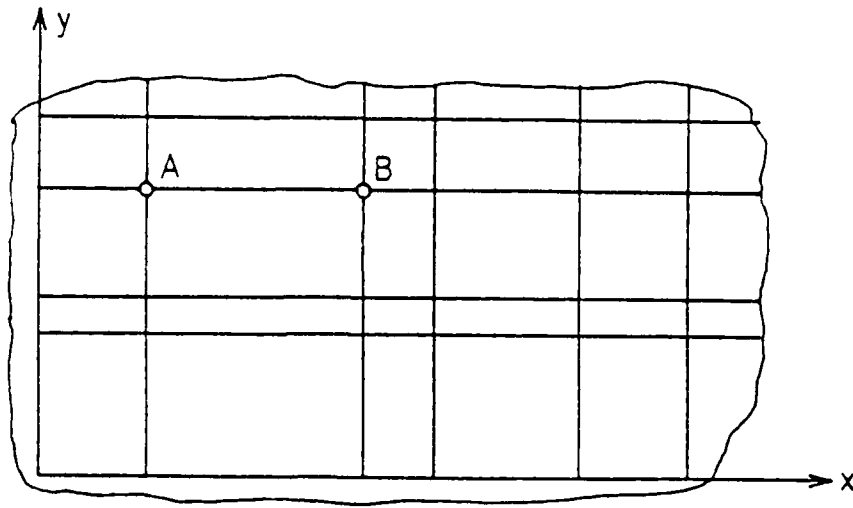


Figure 3.2
Sample element partition in 2 dimensions

Assume that the Hermite nodal values are prescribed at the corners. Then a function may be interpolated into any of the elements. It is an n -cubic polynomial there. This function is continuous and has continuous first derivatives everywhere. Such a function is said to belong to the class C^1 .

The proof is easy but not entirely obvious. Let us sketch it for $n = 2$. Take any segment separating two elements, e. g. the line segment A-B in figure 3.2. The limits of $f(x,y)$ from the top and bottom elements are two cubic polynomials $f_{TOP}(x)$, $f_{BOTTOM}(x)$ in x . However such polynomials are completely determined by the values $f(A)$, $f_x(A)$, $f(B)$, $f_x(B)$ which are common to both polynomials. Hence the polynomials must coincide. The reader may wish to extend the argument to the continuity of f_y across the segment A-B. It becomes transparent why the mixed derivatives f_{xy} are needed at the nodes.

3.1.6. Compound elements.

We will encounter functions which are rough in some areas and smooth in others. In order to represent them properly and economically we like to be able to change the size of the elements in a way that is more flexible than that indicated in figure 3.2. In order to achieve this, we must sacrifice something, namely the simplicity of the elements. It will be sufficient to outline the procedure in R^2 and to consider the "compound element" shown in figure 3.3

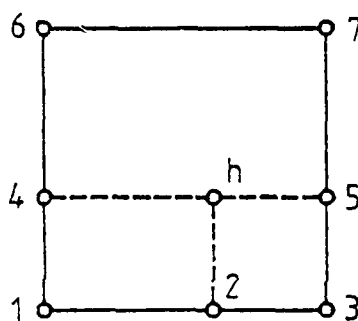


Figure 3.3. Sample of a compound element

The compound element shown has $n = 7$ nodes. At each node i four nodal parameters are prescribed. They are $f(i)$, $f_y(i)$, $f_x(i)$, $f_{xy}(i)$. If a point u is to be interpolated which is situated in the upper quad, we just take formula (3.13) specified above. Note that also the artificial node h may be viewed as a node of the upper quad. Its nodal values $f(h)$, $f_y(h)$, $f_x(h)$, $f_{xy}(h)$ are interpolated linearly from those of $i = 4$ and $i = 5$ alone and do not depend on those of $i = 6$, $i = 7$. Having nodal parameters in h , it poses no difficulties to interpolate points in the lower quads. The interpolation formula for any point x in the compound quad may now be written as:

$$f(x) = \sum_{i=1}^7 \sum_{j=1}^4 c_{ij} \sigma_{ij}(x, y) \quad (3.15)$$

The index i refers now to the nodes, the index j to the nodal derivative. The basis functions $\sigma_{ij}(x,y)$ of the compound quad are piecewise cubic polynomials. They are composed of cubic polynomials having as domain the 3 subquads making up the compound quad. Note that the artificial node h does not enter the interpolation formula. Its nodal parameters are linear functions of those at the genuine nodes, and have thus been eliminated. Formula (3.15) holds also for simple quads, in which case the $\sigma_{ij}(x,y)$ are just the $\gamma_{i_x i_y j_x j_y}(x,y)$ in a different notation.

C^1 continuity within the compound quad is obvious. It is further obvious that C^1 continuity holds within a region partitioned into simple and compound quads in a way that any node of a quad is shared by all neighbouring quads. Figure 3.4 gives an example of such a region.

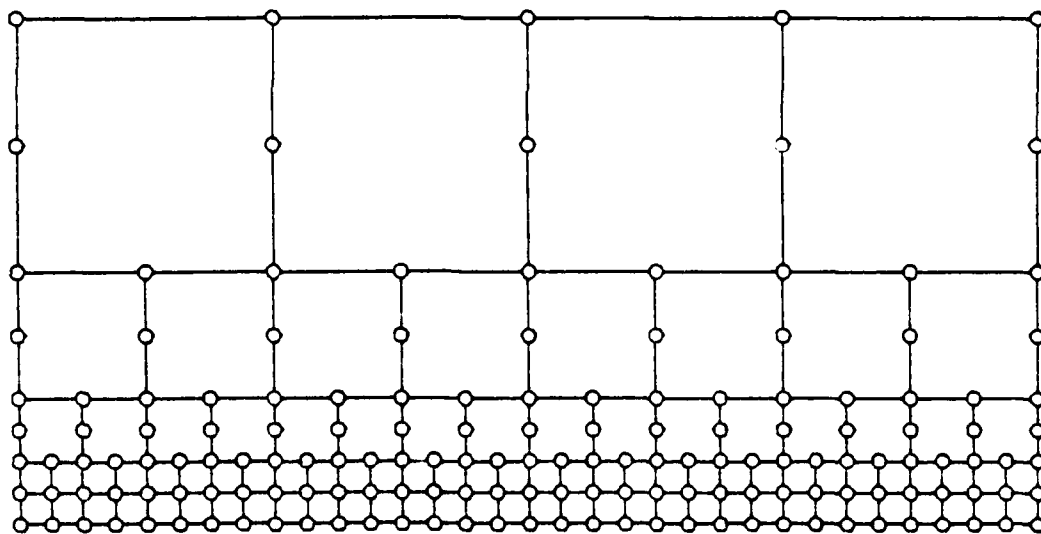


Figure 3.4. Element partition in R^2 using simple and compound elements

Of course, alternative shapes of compound elements can be designed.

Figure 3.5 shows some of them.

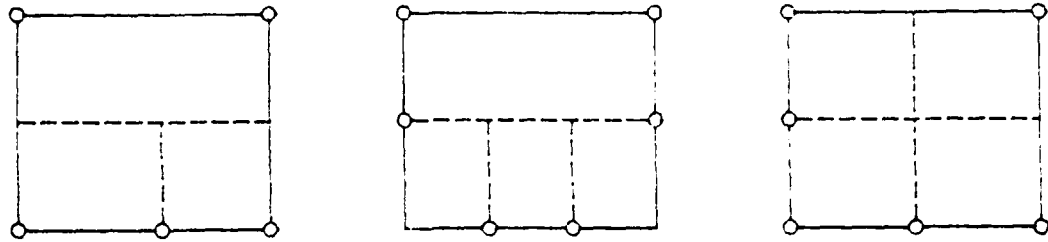


Figure 3.5
Other examples of compound elements

The idea is always the same: decompose the elements into subelements of simple, i.e. rectangular shape. Make a choice of nodes which you like to retain in the final compound element. The nodal parameters at the other nodes must be linearly expressible in terms of the parameters of the retained nodes by linear interpolation (not extrapolation!). After specifying nodal parameters at the retained nodes, interpolation of any location is done by a formula like (3.15). The basis functions $\sigma(x,y)$ are obtained by specifying parameter values equal to zero except for one parameter where a value of 1 is specified.

3.2. Shape functions.

For presentational purposes it is useful to introduce shape functions. Consider a region subdivided into finite elements as for example that one shown in figure 3.4. Label the nodes in some way by $i = 1, 2, \dots$. At each node, label the nodal parameters (e.g. f , f_y , f_x , f_{xy}) by $i = 1, 2, 3, 4$. For each pair (i, j) consider a function $S_{ij}(x, y)$ which has zero nodal values at all nodes $i' \neq i$. Furthermore, at node i , all nodal parameters $j' \neq j$ are also zero. $S_{ij}(x, y)$ will be nonzero only in elements adjacent to the node i . This feature is a great advantage

because precisely the locality of the shape functions $S_{ij}(x,y)$ is responsible for the sparsity of the normal equations to be derived later. In a particular element adjacent to node i , $S_{ij}(x,y)$ of course coincides with the local functions $\sigma_{ij}(x,y)$ introduced earlier (cf. equ. (3.15)). Hence the shape functions $S_{ij}(x,y)$ are nothing new. In the later developments we will mainly deal with their fragments, the element related $\sigma_{ij}(x,y)$'s. However, many interrelationships are more clearly explained, if the globally defined functions $S_{ij}(x,y)$ are used.

3.3. Contribution of the field to the normal equations.

In conventional least squares setups, the normal equations are formed from observations. Observations will also contribute to the normal equations in our case. However there will be another contribution. The harmonicity of the field is not automatically implied by the Hermite cubic representation. Complete harmonicity is practically incompatible with this field representation. All that can be done is an approximate fulfillment of Laplace's equation. This will be achieved by minimizing the integral over the squared Laplacean of the field. This integral will give the additional contribution to the normals as announced earlier.

3.3.1. Reasons for excluding the Ritz method.

Our least squares approach may be called an old fashioned one. At least this is indicated in current treatises on finite elements (cf. Strang; Fix (1973), pp. 133 to 134). In these treatises, attention is focused on more modern principles due to Ritz and Trefftz and generalizations of them published in the last few decades. Confer Oden (1979). Let us explain why we do not propose the Ritz principle. (I tried it in numerical experiments, but it gave poor results!). The Ritz principle

is successfully used in the following typical situation:

Find a solution of

$$\begin{aligned} \Delta V &= 0 & \text{in } B \\ V &= g & \text{on } \partial B \end{aligned} \quad (3.16)$$

This is the familiar Dirichlet problem. However not Dirichleticity is the point. We could have used Neumanns boundary conditions as well, or even a mixture of both. The decisive point is that fixed boundary values are prescribed. The Ritz principle replaces the above problem by a variational one:

Find a solution of

$$\frac{1}{2} \int_B |\text{grad } V|^2 dB = \text{Min.} \quad (3.17)$$

subject to

$$V = g \quad \text{on } \partial B \quad (3.18)$$

The variational formula is slightly more complicated in case of Neumanns boundary condition, but this is irrelevant presently.

The next step is to replace V in (3.17) by its finite element representation, i. e. by

$$V = \sum_{ij} V_{ij} S_{ij}(x,y) \quad (3.19)$$

Here V_{ij} are the unknown nodal parameters and $S_{ij}(x,y)$ are the known shape functions. The functional (integral) to be minimized becomes a quadratic function in the unknowns V_{ij} . The fulfillment of the boundary conditions $V = g$ at ∂B can not be postulated in a strict sense. One has to be satisfied that $V = g$ at certain points of the boundary

and perhaps also that some derivatives of V and g along the boundary coincide at these points. The boundary conditions thus yield a linear set of constraints for the unknowns V_{ij} . Minimization of a quadratic functional subject to linear side constraints is a standard problem which leads to the familiar linear normal equations whose solution are the V_{ij} .

Consider a sequence of partitions of B into finite elements such that the diameter of the largest element goes to zero. It is also required that the shape of the elements is not too badly distorted as the diameter tends to zero. In treatises on finite elements it is proved that under fairly general conditions the finite element solution converges to the exact solution of the original problem. The main advantage of the Ritz method over the more primitive method indicated above and to be described in detail below, namely the method of minimizing

$$\frac{1}{2} \int_B (\Delta V)^2 dB \quad (3.20)$$

subject to

$$V = g \text{ at } \partial B \quad (3.21)$$

is the following one: The Ritz method involves a functional defined in terms of first derivatives. As a consequence, one may use shape functions $S_{ij}(x,y)$ which are simpler than those required for the other method. The functional (3.20) is defined in terms of second derivatives. It is shown in the literature, that our piecewise cubic polynomials, which are C^1 continuous across element boundaries, are an admissible set of trial functions for the least squares problem (3.20, 3.21). However the Ritz problem can be treated successfully with trial functions which are only piecewise bilinear and just C^0 continuous across element boundaries. Such functions require only one nodal parameter, namely the

function value f at the node. This holds in 1, 2 or 3 dimensions. Recall that our C^1 continuous shape function require 2^d nodal parameters, i.e. 2, 4 or 8 depending on the dimension d . The decrease in number of parameters offers one advantage although it must be balanced against the need to use smaller elements because of the more primitive nature of the shape functions. In any case, the computer programs turn out to be much simpler if bilinear shape functions are used.

But why does the Ritz method not work in our geodetic environment? In geodesy the boundary values are the result of measurements. Measurements are not performed everywhere and they are subject to observation errors. In some areas the measurements are redundant; e.g. there may be measurements of V as well as of some components of the gradient of V . In other areas the measurements may be sparse. There may also be measurements in the interior of B . The constraints in our minimization problem become weighted constraints, so to speak. A problem of balancing the weights arises. The normal equations are the sum of two contributions, namely that one from the minimization of the functional, and that one from the observation equations. Symbolically

$$(p_1 G_1 + p_2 G_2) x = r \quad (3.22)$$

Suppose that the functional to be minimized is the Ritz-functional, i.e.

$$\frac{1}{2} \int_B |\text{grad } V|^2 dB \quad (3.23)$$

If we choose p_1 large and p_2 small, then the gradient of V will be made small at the cost of large residuals at the observations. In the limit $p_1 \rightarrow \infty$, we get a constant V . If we choose p_2 large in comparison to p_1 , we treat the observation equations practically as constraints. The minimization procedure will try to match the observations exactly. This may lead to absurd results in case of redundant observations. Even in

case of nonredundancy, the resulting function V may be far off from that one minimizing the gradient. Hence harmonicity is not ensured. Another disadvantage is encountered in subsequent error propagation studies. If one tries to propagate observation errors to some desired quantity, e.g. a gradient at satellite altitude, the resulting error will reflect the small p_1 rather than the large p_2 . This is extremely undesirable, because the propagated error is then mainly due to a "poorly observed" gradient rather than to the observation errors. These disadvantages are to a large extent avoided if we choose the minimizing functional as

$$\frac{1}{2} \int_B (\Delta V)^2 dB \quad (3.24)$$

From the above discussion it is clear that weights p_1 should be rather large in comparison to p_2 . p_1 weighs now the failure of V to be harmonic, whereas in the earlier case it weighted the failure of V to be a constant function. There is a lot of harmonic functions which deviate considerably from a constant function.

If we minimize the functional (3.24), there will be deviations of ΔV from the zero function; but these deviations will be small. Hence the weights p_1 can be chosen large in comparison to p_2 . Errors propagated after adjustment will therefore mainly reflect the errors due to the smaller weights p_2 which belong to the observations.

3.3.2. The least squares contribution of the field.

After discussing the reasons for choosing the functional (3.24) for the purpose to ensure approximate harmonicity, we turn to the technicalities of calculating the contribution of this functional to the normal equations. Let us first outline the procedure on hand of the shape functions $S_{ij}(p)$. As explained earlier, the index i refers to a node while the index j refers to one of the parameters at this node. The

argument p may now be viewed as 1, 2 or 3 dimensional. In case of detailed discussions we shall mainly stress the case of 2-dimensional polar coordinates, where we write $S_{ij}(r, \varphi)$ instead of $S_{ij}(p)$.

The potential is represented as

$$V = \sum_{ij} V_{ij} S_{ij}(p) \quad (3.25)$$

Its Laplacean is

$$\Delta V = \sum_{ij} V_{ij} \Delta S_{ij}(p) \quad (3.26)$$

If 2-dimensional polar coordinates are used, the Laplacean is

$$\Delta V = V_{rr} + \frac{1}{r} V_r + \frac{1}{r^2} V_{\varphi\varphi} \quad (3.27)$$

In order to have something specific in mind, the reader may imagine the region B in the form of a circular ring. In polar coordinates, the ring becomes a rectangular region. It may be subdivided into simple and compound elements as shown in figure 3.6(a). The subdivision shown there corresponds to a subdivision of the ring which is shown in figure 3.6(b).

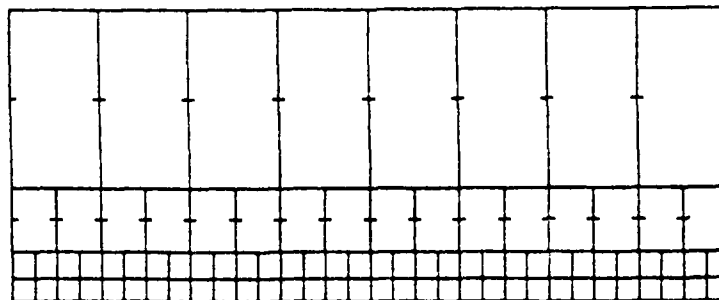


Figure 3.6.(a). Partition of a circular ring represented in polar coordinates

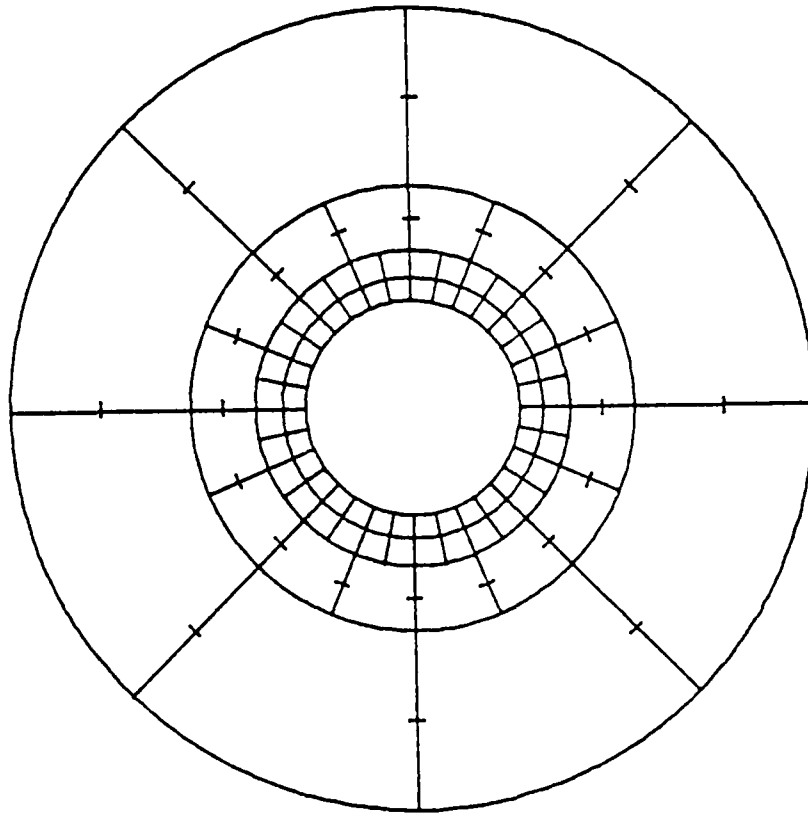


Figure 3.6.(b). Partition of a circular ring represented in cartesian coordinates

Such a subdivision may be chosen if the field is anticipated to be more detailed in the vicinity of the inner boundary. Substituting the Laplacean (3.27) into the functional (3.24) gives

$$E = \frac{1}{2} \int_{\mathcal{B}} \left(\sum_{i,j} V_{ij} \Delta S_{ij}(p) \right)^2 dB(p) \quad (3.28)$$

Minimizing with respect to the V_{ij} leads to the normal equations

$$\frac{\partial E}{\partial V_{ij}} = 0 \quad \begin{array}{l} i \dots \text{nodes} \\ j \dots \text{parameters/node} \end{array} \quad (3.29)$$

or

$$\sum_{i,j'} \left[\int_B \Delta S_{ij}(p) \Delta S_{i'j'}(p) dB(p) \right] V_{i'j'} = 0 \quad (3.30)$$

$i \dots \text{nodes}$
 $j \dots \text{parameters/node}$

Up to a multiplication of these equations by the weight p_i , we get the contribution of the functional (3.24) to the normal equations. Note that for a particular equation labeled (i,j) , there are only a few nonzero coefficients. If we write the equations as

$$\sum_{i,j'} g_{ij;i'j'} V_{i'j'} = 0 \quad (3.31)$$

then

$$g_{ij;i'j'} = 0 \quad (3.32)$$

unless the nodes i and i' are "neighbored". Two nodes are neighbored if there is an element in which both nodes participate. The sparse structure of the normal equations becomes visible now.

3.4. Detailed instructions for computer implementation.

If one works with a computer, the above outlined procedure is not recommended. It is preferable to decompose the total "energy" E

into contributions of the individual elements now denoted B_k .

$$E = \sum_k E_k \quad (3.33)$$

with

$$E_k = \frac{1}{2} \int_{B_k} \left[\sum_{i,j'} V_{i,j'} \Delta S_{i,j'}(p) \right]^2 dB(p) \quad (3.34)$$

In agreement with the earlier notation $\sigma_{ij}(p)$ for the element-internal basis functions, this may be written as

$$E_k = \frac{1}{2} \int_{B_k} \left[\sum_{i,j'} V_{i,j'} \Delta \sigma_{i,j'}(p) \right]^2 dB(p) \quad (3.35)$$

The summation needs to be extended only over nodes participating in the element. We may form the "partial normals" due to the contribution of element k :

$$\sum_{i,j'} \int_{B_k} [\Delta \sigma_{ij}(p) \Delta \sigma_{i,j'}(p) dB(p)] V_{i,j'} = 0 \quad (3.36)$$

or

$$\sum_{i,j'} g_{ij; i,j'}^{(k)} V_{i,j'} = 0 \quad (3.37)$$

If there are I nodes participating in the element, and if J is the number of parameters per node, the above is a symmetric system of IJ equations.

The normal equations (3.31) for the entire region are obtained by summing all partial normal equations. The situation is similar to a network adjustment where partial normals may be formed for subsets

of the measurements; a subset may e.g. be given by a round of directions taken at one station. The partial normals are then combined to yield the normals of the entire network.

3.4.1. The partial normals of a simple quad.

We return to the notation of section 3.1.2. for the basis functions in a simple quad writing:

$$V(r, \varphi) = \sum_{i_r' i_{\varphi}' j_r' j_{\varphi}'} V_{i_r' i_{\varphi}' j_r' j_{\varphi}'} \gamma_{i_r' j_r'}(r-r_0; r_1-r_0) \cdot \gamma_{i_{\varphi}' j_{\varphi}'}(\varphi-\varphi_0; \varphi_1-\varphi_0) \quad (3.38)$$

The simple quad is assumed to extend over the region

$$\begin{aligned} r_0 &\leq r \leq r_1 \\ \varphi_0 &\leq \varphi \leq \varphi_1 \end{aligned} \quad (3.39)$$

Our notation appears complicated now. However we are trying to give detailed instruction for an efficient computer implementation rather than trying to please the casual reader with some slick and polished notation which suppresses the nasty details.

We form the Laplacean (confer (3.9) on the notation for the derivatives of the γ 's)

$$\begin{aligned} \Delta V(r, \varphi) = & \sum_{i_r' i_{\varphi}' j_r' j_{\varphi}'} V_{i_r' i_{\varphi}' j_r' j_{\varphi}'} \left\{ \gamma_{2 i_r' j_r'}(r-r_0; r_1-r_0) \cdot \gamma_{0 i_{\varphi}' j_{\varphi}'}(\varphi-\varphi_0; \varphi_1-\varphi_0) \right. \\ & + \frac{1}{r} \gamma_{1 i_r' j_r'}(r-r_0; r_1-r_0) \cdot \gamma_{0 i_{\varphi}' j_{\varphi}'}(\varphi-\varphi_0; \varphi_1-\varphi_0) + \\ & \left. + \frac{1}{r^2} \gamma_{0 i_r' j_r'}(r-r_0; r_1-r_0) \cdot \gamma_{2 i_{\varphi}' j_{\varphi}'}(\varphi-\varphi_0; \varphi_1-\varphi_0) \right\} \end{aligned} \quad (3.40)$$

Let us abbreviate this for the moment as

$$\Delta V(r, y) = \sum_{i_r i_y j_r j_y} V_{i_r i_y j_r j_y} \sum_{\kappa, \kappa'=1}^3 f_{i_r j_r}^{(\kappa)}(r) h_{i_y j_y}^{(\kappa')}(y) \quad (3.41)$$

In order to form the normals (3.36, 3.37) we have to evaluate all integrals

$$I_{i_r i_y j_r j_y; i_r' i_y' j_r' j_y'}^{KK'} = \int_{r=r_0}^{r_1} \int_{y=y_0}^{y_1} f_{i_r j_r}^{(\kappa)}(r) f_{i_r' j_r'}^{(\kappa')}(r) h_{i_y j_y}^{(\kappa)}(y) h_{i_y' j_y'}^{(\kappa')}(y) r dr dy \quad \dots (3.42)$$

The coefficients of the partial normals

$$\sum_{i_r i_y j_r j_y; i_r' i_y' j_r' j_y'} g_{i_r i_y j_r j_y; i_r' i_y' j_r' j_y'} V_{i_r i_y j_r j_y} = 0 \quad (3.43)$$

are then given by

$$g_{i_r i_y j_r j_y; i_r' i_y' j_r' j_y'} = \sum_{\kappa, \kappa'=1}^3 I_{i_r i_y j_r j_y; i_r' i_y' j_r' j_y'}^{KK'} \quad (3.44)$$

It is a great computational advantage that the above integrals decompose into one-dimensional integrals

$$I_{i_r i_y j_r j_y; i_r' i_y' j_r' j_y'}^{KK'} = I_{i_r j_r i_r' j_r'}^{KK'} \cdot I_{i_y j_y i_y' j_y'}^{KK'} \quad (3.45)$$

with

$$I_{i_r j_r i_r' j_r'}^{KK'} = \int_{r=r_0}^{r_1} f_{i_r j_r}^{(\kappa)}(r) \cdot f_{i_r' j_r'}^{(\kappa')}(r) r dr \quad (3.46)$$

$$I_{i_y j_y i_y' j_y'}^{KK'} = \int_{y=y_0}^{y_1} h_{i_y j_y}^{(\kappa)}(y) \cdot h_{i_y' j_y'}^{(\kappa')}(y) dy \quad (3.47)$$

It is sufficient to outline the further procedure on hand of one of these integrals. Let us take as an example

$$I_{i_r j_r i_r' j_r'}^{13} = \int_{r=r_0}^{r_1} \psi_{2i_r j_r}(r-r_0; r_1-r_0) \psi_{0i_r' j_r'}(r-r_0; r_1-r_0) \frac{dr}{r} \quad (3.48)$$

We substitute

$$r = r_0 + \varrho(r_1 - r_0) \quad (3.49)$$

then

$$I_{i_r j_r i_r' j_r'}^{13} = \int_0^1 \psi_{2i_r j_r}(\varrho(r_1-r_0); r_1-r_0) \psi_{0i_r' j_r'}(\varrho(r_1-r_0); r_1-r_0) \frac{r_1-r_0}{r_0+\varrho(r_1-r_0)} d\varrho \quad (3.50)$$

By (3.10) in section 3.1.2. this further equals:

$$I_{i_r j_r i_r' j_r'}^{13} = (r_1-r_0)^{j_r+j_r'-2-0+1} \int_0^1 \psi_{2i_r j_r}(\varrho) \psi_{0i_r' j_r'}(\varrho) \frac{d\varrho}{r_0+\varrho(r_1-r_0)} \quad (3.51)$$

The functions $\psi_{ki_r j_r}(\varrho)$ are now derivatives of the basis functions for the unit interval. Confer section 3.1.1., equation (3.6). They are cubic polynomials whose coefficients are stored in the array PSC(K,I,J,L) during the computer runs.

According to the outlined rule, we get for any of the integrals

$$I_{i_r j_r i_r' j_r'}^{KK'} \quad (3.52)$$

an expression of the form

$$I_{i_r j_r i_r' j_r'}^{KK'} = (r_1-r_0)^{j_r+j_r'-d_r(K)-d_r(K')+1} \quad (3.53)$$

$$\int_0^1 \psi_{d_r(K) i_r j_r}(\varrho) \psi_{d_r(K') i_r' j_r'}(\varrho) (r_0+\varrho(r_1-r_0))^{e(K)+e(K')+1} d\varrho$$

analogously

$$I_{i_y j_y i_{y'} j_{y'}}^{k k'} = (\varphi_1 - \varphi_0)^{j_y + j_{y'} - d_y(k) - d_{y'}(k') + 1} \cdot \int_0^1 \Psi_{d_y(k) i_y j_y}(\omega) \Psi_{d_{y'}(k') i_{y'} j_{y'}}(\omega) d\omega \quad (3.54)$$

Thereby

$$\omega = \frac{\varphi - \varphi_0}{\varphi_1 - \varphi_0} \quad (3.55)$$

The following table lists the degrees $d_r(k)$, $d_y(k)$ and the exponents $e(k)$

k	d_r	d_y	e
=====			
1	2	0	0
2	1	0	-1
3	0	2	-2

Table 3.2
Listing $d_r(k)$, $d_y(k)$, $e(k)$; $k=1,2,3$

All these integrals are extended over weighted products of the basis functions and their derivatives for the unit interval. These basis functions are cubic polynomials in one variable. The first and second derivatives are quadratic and linear polynomials respectively.

Hence all the integrals can be assembled with the help of the PSC-table and the following integrals over weighted monomials:

$$M_{st}(r_0, r_1) = \int_0^1 x^s (r_0 + (r_1 - r_0)x)^t dx \quad \begin{matrix} 0 \leq s \leq 6 \\ -3 \leq t \leq 1 \end{matrix} \quad (3.56)$$

Summarizing, we obtain the following procedure to evaluate

$$g_{i_r i_y j_r j_y; i_{r'} i_{y'} j_{r'} j_{y'}}$$

for a simple element $r_0 \leq r \leq r_1$; $y_0 \leq y \leq y_1$

[1] Evaluate the weighted integrals over monomials according to (3.56) and tabulate them for $0 \leq s \leq 6$, $-3 \leq t \leq 1$

[2] Using the PSC-tables, whose coefficients are denoted $c_{\kappa \rho d}$ (cf. equation (3.7)), and the table 3.2 for $d(k)$, $e(k)$, evaluate integrals over products of normalized basis functions

$$\begin{aligned} & \int_0^1 \psi_{d_r(\kappa) i_r j_r}(\varphi) \psi_{d_r(\kappa') i_{r'} j_{r'}}(\varphi) [r_0 + (r_1 - r_0)\varphi] d\varphi = \\ & = \sum_{\ell=0}^3 \sum_{\ell'=0}^3 c_{d_r(\kappa) i_r j_r \ell} c_{d_r(\kappa') i_{r'} j_{r'} \ell'} M_{\ell+\ell', e(\kappa)+e(\kappa')+1} = J_{i_r j_r i_{r'} j_{r'}}^{\kappa \kappa'} \\ & \int_0^1 \psi_{d_y(\kappa) i_y j_y}(\omega) \psi_{d_y(\kappa') i_{y'} j_{y'}}(\omega) d\omega = \\ & = \sum_{\ell=0}^3 \sum_{\ell'=0}^3 c_{d_y(\kappa) i_y j_y \ell} c_{d_y(\kappa') i_{y'} j_{y'} \ell'} M_{\ell+\ell', 0} = J_{i_y j_y i_{y'} j_{y'}}^{\kappa \kappa'} \end{aligned} \quad (3.57)$$

[3] According to equations (3.53), (3.54) evaluate:

$$\begin{aligned} I_{i_r j_r i_{r'} j_{r'}}^{\kappa \kappa'} &= (r_1 - r_0)^{j_r + j_{r'} - d_r(\kappa) - d_r(\kappa') + 1} \cdot J_{i_r j_r i_{r'} j_{r'}}^{\kappa \kappa'} \\ I_{i_y j_y i_{y'} j_{y'}}^{\kappa \kappa'} &= (y_1 - y_0)^{j_y + j_{y'} - d_y(\kappa) - d_y(\kappa') + 1} \cdot J_{i_y j_y i_{y'} j_{y'}}^{\kappa \kappa'} \end{aligned} \quad (3.58)$$

$$0 \leq i_r j_r i_{r'} j_{r'}, i_y j_y i_{y'} j_{y'} \leq 1$$

and tabulate these integrals

[4] Assemble

$$g_{i_r i_y j_r j_y; i_r' i_y' j_r' j_y'} = \sum_{\kappa, \kappa'=1}^3 I_{i_r j_r i_r' j_r'}^{\kappa \kappa'} \cdot I_{i_y j_y i_y' j_y'}^{\kappa \kappa'} \quad (3.59)$$

Thus the partial normals for a simple element are evaluated!

Remark: There are some further computational shortcuts. Because of symmetries, some of the

$$I_{i_r j_r i_r' j_r'}^{\kappa \kappa'} \quad ; \quad I_{i_y j_y i_y' j_y'}^{\kappa \kappa'}$$

are the same. Such symmetries could be exploited, but we have refrained from doing so in our numerical experiments.

Remark: The multi-index-notation is convenient for calculating the coefficients $g_{i_r i_y j_r j_y; i_r' i_y' j_r' j_y'}$. For subsequent calculations it is preferable to switch to a two-dimensional array g_{ij} . The multi-indexes i_r, i_y imply a node numbering $n = 1, 2, 3, 4$ as shown in figure 3.7 and given by

$$n = 2i_r + i_y + 1 \quad (3.60)$$

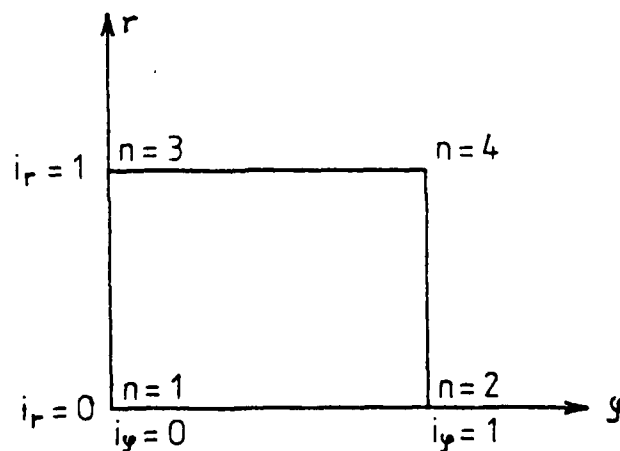


Figure 3.7. Numbering of nodes in a simple quad

The parameter numbering per node has been explained above repeatedly. It is given by

$$q = 2j_r + j_y + 1 \quad (3.61)$$

and shown in table 3.3

q	j _r	j _y	Math. symbol
1	0	0	V
2	0	1	V _y
3	1	0	V _r
4	1	1	V _{ry}

Table 3.3. Parameter numbering at a node.

We adopt an accumulative parameter numbering by

$$i = 4(n - 1) + q \quad (3.62)$$

In this way we get a two-dimensional array by:

$$g_{ii'} = g_{i_r i_y j_r j_y; i'_r i'_y j'_r j'_y} \quad (3.63)$$

with

$$\begin{aligned} i &= 8i_r + 4i_y + 2j_r + j_y + 1 \\ i' &= 8i'_r + 4i'_y + 2j'_r + j'_y + 1 \end{aligned} \quad (3.64)$$

In this way we get a 16x16 matrix for the present example. There are 16 parameters involved in the simple quad, 4 at each of the 4 nodes. The numbering groups the parameters for one node together. Table 3.4 explains all the details once more.

	node	i_r	i_φ	j_r	j_φ	common math. language
1	1	0	0	0	0	V at lower left corner
2	1	0	0	0	1	V_φ --
3	1	0	0	1	0	V_r --
4	1	0	0	1	1	$V_{r\varphi}$ --
5	2	0	1	0	0	V at lower right corner
6	2	0	1	0	1	V_φ --
7	2	0	1	1	0	V_r --
8	2	0	1	1	1	$V_{r\varphi}$ --
9	3	1	0	0	0	V at upper left corner
10	3	1	0	0	1	V_φ --
11	3	1	0	1	0	V_r --
12	3	1	0	1	1	$V_{r\varphi}$ --
13	4	1	1	0	0	V at upper right corner
14	4	1	1	0	1	V_φ --
15	4	1	1	1	0	V_r --
16	4	1	1	1	1	$V_{r\varphi}$ --

Table 3.4. Numbering of nodes and parameters in a simple quad.

This all is rather irrelevant to the casual reader. It is, however, essential to the person implementing the method efficiently on a computer.

Remark: The coefficients g_{ij} are functions of: $r_0, r_1, \varphi_1 - \varphi_0$. It follows that the coefficients are the same for two quads sharing these parameters. If all simple quads in a "layer" $r_0 \leq r \leq r_1$ have the same angular width $\varphi_1 - \varphi_0$, then the coefficients need to be evaluated only once!

3.4.2. Partial normals of a compound element

Again it will be sufficient to consider the two-dimensional case and to describe the procedure on hand of a specific example. It is the compound element described above in section 3.1.6. and shown in figure 3.8

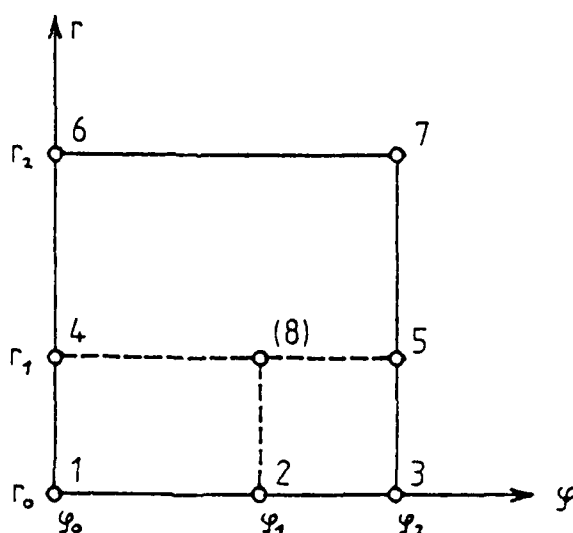


Figure 3.8. Compound quad.

We start by interpolating the parameters at the auxiliary node 8 from those at nodes 4 and 5. The function $V(r, \varphi)$ is given in the upper quad 4, 5, 6, 7 by

$$V(r, \varphi) = \sum_{i=4}^7 \sum_{j=1}^4 V_{ij} \sigma_{ij}(r-r_1, \varphi-\varphi_1) \quad (3.65)$$

It is useful to note that the basis functions belonging to the upper nodes, i.e. σ_{6j}, σ_{7j} vanish together with their first derivatives at the line from 4 to 5. Hence, as long as we are only interested in zero-th and first derivatives of $V(r, \varphi)$ at this line, we may use the simplified

representation:

$$V(r, y) = \sum_{i=1}^5 \sum_{j=1}^4 V_{ij} \sigma_{ij}(r-r_1, y-y_1) + O((r-r_1)^2) \quad (3.66)$$

Using a more explicit representation we may write

$$V(r, y) = \sum_{j_r, j_y=0}^1 \left[V_{4j_r j_y} \gamma_{0j_r}(r-r_1; r_2-r_1) \cdot \gamma_{0j_y}(y-y_1; y_2-y_1) + \right. \\ \left. + V_{5j_r j_y} \gamma_{1j_r}(r-r_1; r_2-r_1) \cdot \gamma_{1j_y}(y-y_1; y_2-y_1) \right] \quad (3.67)$$

The functions $\gamma_{ij}(x; h)$ are defined in section 3.1.2., equation (3.8). We are now able to interpolate the nodal parameters of the auxiliary node 8:

$$V_{8j_r' j_y'} = \sum_{j_r, j_y} \left[V_{4j_r j_y} \gamma_{j_r' 0j_r}(0; r_2-r_1) \gamma_{j_y' 0j_y}(y_2-y_1; y_2-y_1) + \right. \\ \left. + V_{5j_r j_y} \gamma_{j_r' 1j_r}(0; r_2-r_1) \gamma_{j_y' 1j_y}(y_2-y_1; y_2-y_1) \right] \quad (3.68)$$

For the functions $\gamma_{ij}(x)$ see equations (3.9), (3.10). Switching to the more condensed notation

$$j = 2j_r + j_y + 1 \\ j' = 2j_r' + j_y' + 1 \quad (3.69)$$

we write this relationship as

$$V_{8j'} = \sum_{j=1}^4 (u_{jj'} V_{4j} + v_{jj'} V_{5j}) \quad (3.70)$$

In an even more condensed form we write

$$V_8 = (U_4 \ U_5) \begin{pmatrix} V_4 \\ V_5 \end{pmatrix} = U \begin{pmatrix} V_4 \\ V_5 \end{pmatrix} \quad (3.71)$$

where U is an 8x4 matrix.

The next step is to form according to section 3.4.1. dealing with simple quads the partial normals for the lower left subelement. Let these equations be written as

$$\begin{bmatrix} G_{11}^{LL} & G_{12}^{LL} & G_{14}^{LL} & G_{18}^{LL} \\ G_{21}^{LL} & G_{22}^{LL} & G_{24}^{LL} & G_{28}^{LL} \\ G_{41}^{LL} & G_{42}^{LL} & G_{44}^{LL} & G_{48}^{LL} \\ G_{81}^{LL} & G_{82}^{LL} & G_{84}^{LL} & G_{88}^{LL} \end{bmatrix} \cdot \begin{bmatrix} V_1 \\ V_2 \\ V_4 \\ V_8 \end{bmatrix} = 0 \quad (3.72)$$

One eliminates the auxiliary node by the substitution

$$V_8 = (U_4 \ U_5) \begin{pmatrix} V_4 \\ V_5 \end{pmatrix} \quad (3.73)$$

Similar to a parameter transformation in adjustment by variation of parameters, we must consider the full transformation matrix implied by

$$\begin{bmatrix} V_1 \\ V_2 \\ V_4 \\ V_8 \end{bmatrix} = \begin{bmatrix} I & & & \\ & I & & \\ & & I & \\ & & & U_4 \ U_5 \end{bmatrix} \cdot \begin{bmatrix} V_1 \\ V_2 \\ V_4 \\ V_5 \end{bmatrix} \quad (3.74)$$

i. e.

$$V = A \cdot \bar{V}$$

and transform the above normals (3.72), now written as

$$GV = 0 \quad (3.75)$$

according to

$$A^T G A \bar{V} = 0 \quad (3.76)$$

The result is

$$\begin{bmatrix} \bar{G}_{11}^u & \bar{G}_{12}^u & \bar{G}_{14}^u & \bar{G}_{15}^u \\ \bar{G}_{21}^u & \bar{G}_{22}^u & \bar{G}_{24}^u & \bar{G}_{25}^u \\ \bar{G}_{41}^u & \bar{G}_{42}^u & \bar{G}_{44}^u & \bar{G}_{45}^u \\ \bar{G}_{51}^u & \bar{G}_{52}^u & \bar{G}_{54}^u & \bar{G}_{55}^u \end{bmatrix} \cdot \begin{bmatrix} V_1 \\ V_2 \\ V_4 \\ V_5 \end{bmatrix} = 0 \quad (3.77)$$

It is seen that

$$\begin{aligned} \bar{G}_{11}^u &= G_{11}^u \\ \bar{G}_{12}^u &= G_{12}^u \\ \bar{G}_{14}^u &= G_{14}^u + G_{18}^u U_4 \\ \bar{G}_{15}^u &= G_{18}^u U_5 \\ \bar{G}_{22}^u &= G_{22}^u \end{aligned} \quad (3.78)$$

$$\begin{aligned}
 \bar{G}_{24}^{LL} &= G_{24}^{LL} + G_{28}^{LL} U_4 \\
 \bar{G}_{28}^{LL} &= G_{28}^{LL} U_5 \\
 \bar{G}_{44}^{LL} &= G_{44}^{LL} + G_{48}^{LL} U_4 + U_4^T G_{84}^{LL} + U_4^T G_{88}^{LL} U_4 \\
 \bar{G}_{48}^{LL} &= G_{48}^{LL} U_5 + U_4^T G_{88}^{LL} U_5 \\
 \bar{G}_{88}^{LL} &= U_5^T G_{88}^{LL} U_5
 \end{aligned} \tag{3.78}$$

In a similar way we obtain for the lower right quad

$$\begin{bmatrix} \bar{G}_{22}^{LR} & \bar{G}_{23}^{LR} & \bar{G}_{24}^{LR} & \bar{G}_{25}^{LR} \\ \bar{G}_{32}^{LR} & \bar{G}_{33}^{LR} & \bar{G}_{34}^{LR} & \bar{G}_{35}^{LR} \\ \bar{G}_{42}^{LR} & \bar{G}_{43}^{LR} & \bar{G}_{44}^{LR} & \bar{G}_{45}^{LR} \\ \bar{G}_{52}^{LR} & \bar{G}_{53}^{LR} & \bar{G}_{54}^{LR} & \bar{G}_{55}^{LR} \end{bmatrix} \cdot \begin{bmatrix} V_2 \\ V_3 \\ V_4 \\ V_5 \end{bmatrix} \tag{3.79}$$

By a direct application of the simple quad formulas we obtain for the upper element

$$\begin{bmatrix} G_{44}^U & G_{45}^U & G_{46}^U & G_{47}^U \\ G_{54}^U & G_{55}^U & G_{56}^U & G_{57}^U \\ G_{64}^U & G_{65}^U & G_{66}^U & G_{67}^U \\ G_{74}^U & G_{75}^U & G_{76}^U & G_{77}^U \end{bmatrix} \cdot \begin{bmatrix} V_4 \\ V_5 \\ V_6 \\ V_7 \end{bmatrix} \tag{3.80}$$

Now the partial normals (3.77), (3.79), (3.80) of the 3 subelements are added into the normals of the compound quads

$$\begin{bmatrix} G_{11} & G_{12} & G_{13} & G_{14} & G_{15} & G_{16} & G_{17} \\ G_{21} & G_{22} & G_{23} & G_{24} & G_{25} & G_{26} & G_{27} \\ G_{31} & G_{32} & G_{33} & G_{34} & G_{35} & G_{36} & G_{37} \\ G_{41} & G_{42} & G_{43} & G_{44} & G_{45} & G_{46} & G_{47} \\ G_{51} & G_{52} & G_{53} & G_{54} & G_{55} & G_{56} & G_{57} \\ G_{61} & G_{62} & G_{63} & G_{64} & G_{65} & G_{66} & G_{67} \\ G_{71} & G_{72} & G_{73} & G_{74} & G_{75} & G_{76} & G_{77} \end{bmatrix} \cdot \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \\ V_6 \\ V_7 \end{bmatrix} \quad (3.81)$$

Any of the submatrices G_{ij} is a sum

$$G_{ij} = \bar{G}_{ij}^{LL} + \bar{G}_{ij}^{LR} + G_{ij}^U \quad (3.82)$$

where $\bar{G}_{ij}^{LL} = 0$, $\bar{G}_{ij}^{LR} = 0$, $G_{ij}^U = 0$ is assumed, if one of the nodes i or j does not participate in the subquad LL , LR or U .

3.5. More general basis functions.

This section may serve as an outline for the construction of shape functions which (1) have a more general analytical representation than piecewise n -cubic polynomials, (2) are C^1 continuous, and (3) are computationally still reasonably efficient.

The motivation to look for alternative shape functions may come from the desire to use functions which conform to the potentials attenuation in a better way. It appears tempting to replace polynomials in r by polynomials in $1/r$, for example. This looks particularly attractive for elements further away from the earth surface. It also opens a way to represent the potential in the entire exterior of some sphere by a finite number of elements of infinite size. This aspect will

be treated further in the next section.

The computational efficiency of cubic polynomials comes from 2 sources. (1) Such polynomials are easily evaluated, (2) the integrals to be evaluated in order to form the partial normals of a simple element decompose into a moderate number of products of integrals over a single variable. The second feature extends to the basis functions to be outlined below, the first one extends only partly.

As in section 3.1.1., one starts again with the unit interval. One introduces functions $\psi_j(x)$ such that the j -th derivative of $\psi_j(x)$ at the i -th-interval end is unity, while all other derivatives at this and the other interval end vanish. It is important to note that the highest derivative considered this way need not be the same at both interval ends. To have something specific in mind, consider the basis functions

$$\begin{aligned}\psi_{00}(x) &= 3x^4 - 4x^3 && + 1 \\ \psi_{01}(x) &= 2x^4 - 3x^3 && + x \\ \psi_{02}(x) &= \frac{1}{2}x^4 - x^3 + \frac{1}{2}x^2 && (3.83) \\ \psi_{10}(x) &= -3x^4 + 4x^3 \\ \psi_{11}(x) &= x^4 - x^3\end{aligned}$$

In this example the second derivative is only modelled at the left end of the interval, but not at the right end. If the above basis functions

are changed to

$$\begin{aligned}
 \gamma_{00}(x) &= -6x^5 + 15x^4 - 10x^3 + 1 \\
 \gamma_{01}(x) &= -3x^5 + 8x^4 - 6x^3 + x \\
 \gamma_{02}(x) &= -\frac{1}{2}x^5 + \frac{3}{2}x^4 - \frac{3}{2}x^3 + \frac{1}{2}x^2 \\
 \gamma_{10}(x) &= 6x^5 - 15x^4 + 10x^3 \\
 \gamma_{11}(x) &= -3x^5 + 7x^4 - 4x^3 \\
 \gamma_{12}(x) &= \frac{1}{2}x^5 - x^4 + \frac{1}{2}x^3
 \end{aligned} \tag{3.84}$$

then we are dealing with Hermite quintic polynomials. At each node the zero, first and second derivative can be prescribed.

With the chosen set of basis functions for the unit interval one can proceed to any finite interval by (cf. section 3.1.2):

$$\gamma_{ij}(x;h) = h^j \gamma_{ij}\left(\frac{x}{h}\right) \tag{3.85}$$

As in section 3.1.3., two dimensional basis functions are again obtained by forming products

$$\gamma_{i_x i_y j_x j_y}(x,y) = \gamma_{ij}(x;a) \cdot \bar{\gamma}_{ij}(y;b) \tag{3.86}$$

As indicated by the bar over the last γ , there is no need to have the same type of basis functions for x and for y . For example, one may use quintics in x and cubics in y .

In order to ensure C^1 continuity, the type of basis functions used in the x -direction should be matched by the element adjacent in the y direction and vice versa. This poses some problems if one fuses simple elements to compound elements. However, as long as the basis functions are polynomials, there is normally a way out of any dilemma. In any

case, C^1 continuity must be checked very carefully in all unusual situations.

Remark: If quintics are used consistently, the resulting function is even C^2 continuous. This may look attractive. Among others, it offers a way to force the Laplacean to be zero at all nodes. However, a heavy price has to be paid in terms of d^3 parameters per node (d ... number of dimensions). For $d = 3$ there are 27 parameters per node, while in case of bicubics there are only 8.

The nodal parameters needed in these more general finite element representations follow automatically from the one-dimensional basis functions used. Two dimensions are typical enough for our outline here. Any node has a degree in the x and y -direction. These two degrees need not be equal, but they pose restrictions to the basis functions of adjacent elements. Suppose the degree is 2 in the x -direction and 1 in the y -direction. This holds if quintics in x are multiplied with cubics in y . Then the 2×3 nodal parameters are automatically given by all symbolic products of any two elements out of the sets

$$\left\{ 1, \frac{\partial}{\partial x}, \frac{\partial^2}{\partial x^2} \right\}, \quad \left\{ 1, \frac{\partial}{\partial y} \right\} \quad (3.87)$$

i.e. they are:

$$V, V_y, V_x, V_{xy}, V_{xx}, V_{xy} \quad (3.88)$$

The integration procedure outlined in section 3.4.1. and the partial normals of a simple element must be modified if other basis functions than cubic polynomials are used. The modifications are, however, moderate. They amount to a change of the integration formulas for the required products of one-dimensional basis functions and their

derivatives. The factors of these products were denoted by

$$I_{i_r j_r i_r' j_r'}^{KK'} ; I_{i_y j_y i_y' j_y'}^{KK'} \quad (3.89)$$

Confer section 3.4.1, in particular equation (3.59).

3.6. Elements extending to infinity in one direction.

The case of an element extending to infinity in one direction requires special attention. We shall deal with this problem in detail, because use of such elements is to be made in our computer experiments. Consider an infinite interval $r_0 \leq r < \infty$. We choose basis functions

$$\begin{aligned} \chi_{00}(r) &= \frac{r_0^2}{r^3} (3r - 2r_0) \\ \chi_{01}(r) &= \frac{r_0^3}{r^3} (r - r_0) \end{aligned} \quad (3.90)$$

They tend to zero as r^{-2} for $r \rightarrow \infty$. Figure 3.9 shows the graphs of these basis functions.

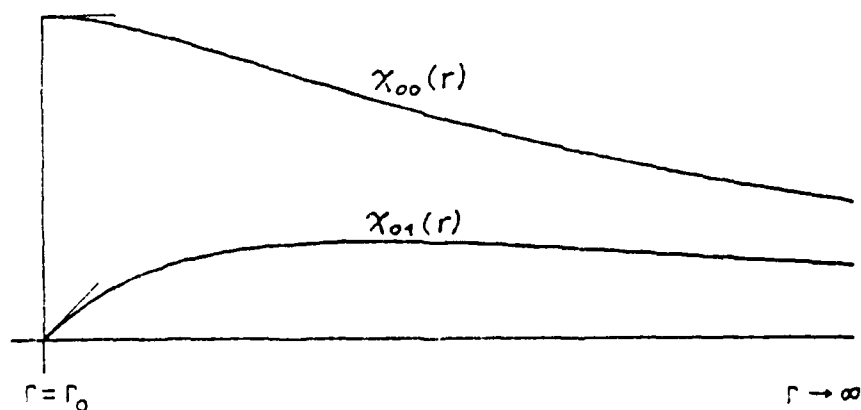


Figure 3.9. Sample basis functions for an interval extending to infinity

We combine them e.g. with cubics in y , which we denote by

$$\psi_{ij}(y-y_0; y_1-y_0) \quad y_0 \leq y \leq y_1 \quad (3.91)$$

Hence we deal with the basis functions

$$\psi_{i_x i_y j_x j_y}(r; y-y_0) = \chi_{i_x j_x}(r) \psi_{i_y j_y}(y-y_0; y_1-y_0) \quad (3.92)$$

$$i_x = 0; \quad j_x, i_y, j_y = 0, 1$$

There is no C^1 continuity problem if an element partition is chosen as shown in figure 3.10 and if bicubics are used for all elements of finite size.

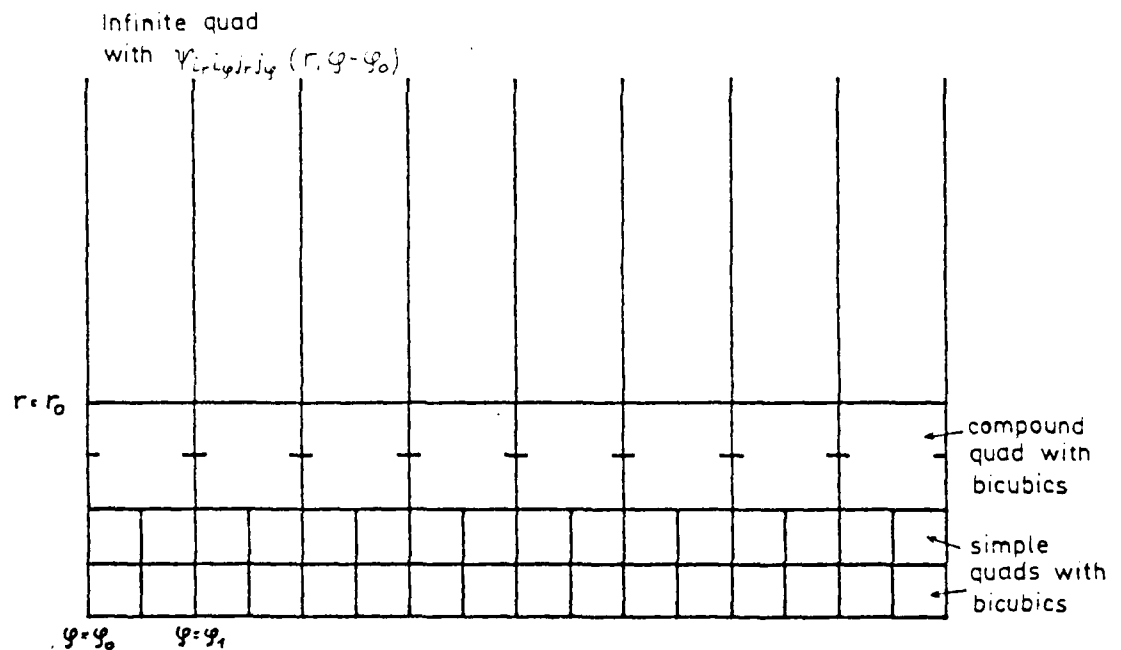


Figure 3.10. Element partition making use of infinite elements

Next we outline the integration procedure in detail for the infinite element. Confer the parallel developments of section 3.4.1. on the partial normals of a simple quad. The representation of the potential $V(r, \varphi)$ in a s r, $\varphi_0 \leq \varphi \leq \varphi_1$ is

$$V(r, \varphi) = \sum_{i_r' i_{\varphi}' j_r' j_{\varphi}'} V_{i_r' i_{\varphi}' j_r' j_{\varphi}'} \chi_{i_r' j_r'}(r) \gamma_{i_{\varphi}' j_{\varphi}'}(\varphi - \varphi_0; \varphi_1 - \varphi_0) \quad (3.93)$$

The index i' is fixed at zero (we retain it only for ease of comparison with section 3.4.1., equation (3.38)). All other indices run from 0 to 1. The Laplacean is:

$$\begin{aligned} \Delta V(r, \varphi) = & \sum_{i_r' i_{\varphi}' j_r' j_{\varphi}'} V_{i_r' i_{\varphi}' j_r' j_{\varphi}'} \left\{ \chi_{2 i_r' j_r'}(r) \gamma_{0 i_{\varphi}' j_{\varphi}'}(\varphi - \varphi_0; \varphi_1 - \varphi_0) + \right. \\ & + \frac{1}{r} \chi_{1 i_r' j_r'}(r) \gamma_{0 i_{\varphi}' j_{\varphi}'}(\varphi - \varphi_0; \varphi_1 - \varphi_0) + \\ & \left. + \frac{1}{r^2} \chi_{0 i_r' j_r'}(r) \gamma_{2 i_{\varphi}' j_{\varphi}'}(\varphi - \varphi_0; \varphi_1 - \varphi_0) \right\} \end{aligned} \quad (3.94)$$

This equation is abbreviated as

$$\Delta V(r, \varphi) = \sum_{i_r' i_{\varphi}' j_r' j_{\varphi}'} V_{i_r' i_{\varphi}' j_r' j_{\varphi}'} \sum_{k=1}^3 f_{i_r' j_r'}^{(k)}(r) \cdot h_{i_{\varphi}' j_{\varphi}'}^{(k)}(\varphi) \quad (3.95)$$

The principle

$$\frac{1}{2} \int_{r=r_0}^{\infty} \int_{\varphi=\varphi_0}^{\varphi_1} (\Delta V(r, \varphi))^2 r dr d\varphi = \text{Min} \quad (3.96)$$

leads to the partial normals of the infinite quad:

$$\sum g_{i_r' i_{\varphi}' j_r' j_{\varphi}'; i_r' i_{\varphi}' j_r' j_{\varphi}'} V_{i_r' i_{\varphi}' j_r' j_{\varphi}'} = 0 \quad (3.97)$$

In analogy to section 3.4.1., the coefficients follow from the formula:

$$g_{i_r i_y j_r j_y i_r' i_y' j_r' j_y'} = \sum_{\kappa, \kappa'=1}^3 I_{i_r j_r i_r' j_r'}^{\kappa \kappa'} \cdot I_{i_y j_y i_y' j_y'}^{\kappa \kappa'} \quad (3.98)$$

The integrals

$$I_{i_y j_y i_y' j_y'}^{\kappa \kappa'}$$

are unchanged. On the other hand:

$$I_{i_r j_r i_r' j_r'}^{\kappa \kappa'} = \int_{r_0}^{\infty} \chi_{d_r(\kappa) i_r j_r}(r) \cdot \chi_{d_r(\kappa') i_r' j_r'}(r) \cdot r^{e(\kappa)+e(\kappa')+1} dr \quad (3.99)$$

Table 3.3 of section 3.4.1. for the d_r , e is still valid. If one writes the $\chi_{d_r i_r j_r}$ in the form

$$\chi_{d_r i_r j_r}(r) = \sum_{\ell=0}^1 c_{d_r i_r j_r \ell} \frac{r_0^{2+j_r+\ell}}{r^{2+d_r+\ell}} \quad (3.100)$$

then a compact 3x1x2x2 table of the coefficients

$$c_{d_r i_r j_r \ell} \quad \begin{array}{l} 0 \leq d_r \leq 2 \\ i_r = 0 \\ 0 \leq j_r \leq 1 \\ 0 \leq \ell \leq 1 \end{array} \quad (3.101)$$

may be stored in the computer. This CHC-Tab is the counterpart to the PSC-Table introduced in section 3.1.1. With the help of the CHC-Table one computes:

$$I_{i_r j_r i_y j_y i_r' i_y' j_r' j_y'}^{\kappa \kappa'} = \sum_{\ell=0}^1 \sum_{\ell'=0}^1 c_{d_r(\kappa) i_r j_r \ell} c_{d_r(\kappa') i_r' j_r' \ell'} \cdot \frac{1}{p_{\kappa \kappa' \ell \ell'} - 1} r_0^{j_r + j_r' - d_r(\kappa) - d_r(\kappa') + e(\kappa) + e(\kappa') + 2} \quad (3.102)$$

Thereby

$$P_{\kappa\kappa'\ell\ell'} = 2 + d_r(\kappa) + \ell + 2 + d_r(\kappa') + \ell' - e(\kappa) - e(\kappa') - 1 \quad (3.103)$$

3.7. The outer zone.

The infinite outer space of the earth can not be partitioned into infinitely many elements of finite size. The field is most detailed in the vicinity of the surface. Hence fairly small elements must be chosen there. The elements may increase as we go outward. The increase in size is partly achieved by the natural increase of the surface element $rdrd\varphi$ or $r^2 \cos \varphi dr d\varphi d\lambda$ in polar coordinates. To a greater extent it is achieved by a lumping of elements as discussed in the section on compound elements. At a certain level the finite element partition may stop at all. The field is then represented in a spherical shell (circular ring). It is forced to be (nearly) harmonic there by minimizing the integral over squared Laplacean. It is now necessary to ensure in some way that the field in the outer space of the shell is

- (1) consistent with the field in the interior of the shell,
- (2) approaching zero if the radius tends to zero, and finally
- (3) harmonic.

Consistency means in a strict sense, that the outer field is the analytical continuation of the field within the shell. In a finite element context this requirement must be somewhat relaxed. Not even within the shell do we have a strictly analytical field. However, if we have C^1 continuity there, as we do when cubic polynomials are used, approximate C^1 continuity across the outer boundary of the shell is a minimum requirement.

There are various ways to deal with the problem of the outer

field. We shall outline them now.

(1) Choose the radius of the outer shell large enough and force the field and its first derivatives to be zero there. A large enough radius means that the size of field and its derivatives must be below the desired computational accuracy. It must be borne in mind that our field is actually a disturbing field superimposed upon a reference field.

(2) Partition of the outer space into finitely many elements of infinite size. Refer to the previous section 3.6 on elements extending to infinity in one direction.

(3) Use of Greens formula. If potential and (outer) normal derivative are prescribed at the boundary Γ , then the potential in the exterior of Γ can be represented by Greens 3rd identity. (See e.g. Heiskanen-Moritz (1967), equation 1-29').

$$4\pi V(P) = \int_{\Gamma} \left\{ V(Q) \frac{\partial}{\partial n} \left(\frac{1}{l(P,Q)} \right) - \frac{\partial V(Q)}{\partial n} \cdot \frac{1}{l(P,Q)} \right\} d\Gamma(Q) \quad (3.104)$$

This formula refers to the 3-dimensional case, $l(P,Q)$ is the distance between the reference point P and the point of integration Q . Confer figure 3.11

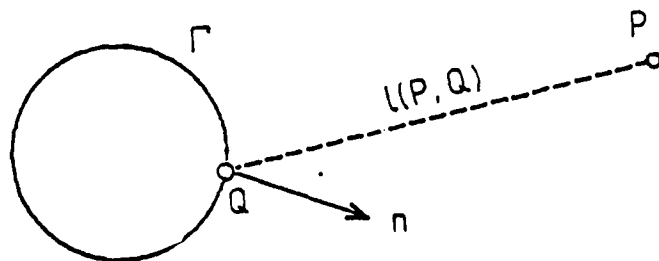


Figure 3.11. Explains the definition of outer normal n and distance l

Consider now an element partition within a shell around the earth surface as shown in figure 3.12.

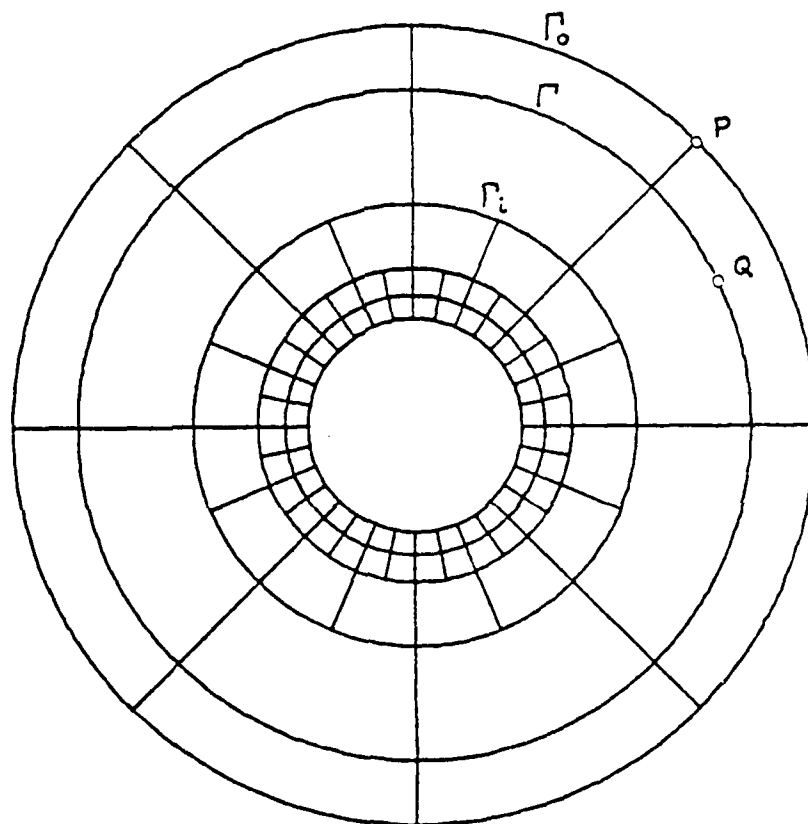


Figure 3.12. Fusing the inner and the outer zone

The surface Γ of integration in equ. (3.104) is assumed spherical (circular) and situated slightly inside the outer boundary Γ_0 of the shell. Greens third identity may now be applied to any node P situated at the outer boundary denoted Γ_0 . Since $V(Q)$ and $\frac{\partial V(Q)}{\partial n}$ is expressed linearly in terms of the parameters belonging to nodes situated in the layer between Γ_0 and Γ_i , we get linear equations relating $V(P)$ to nodal parameters of the outer layer (bounded by Γ_i and Γ_0). Let $V(Q)$ be

represented as

$$V(Q) = \sum_{i \in L} \sum_j V_{ij} S_{ij}(Q) \quad (3.105)$$

Where L represents the set of nodes participating in the outer layer, and j runs over the parameters of any node. The $S_{ij}(Q)$ are the shape functions introduced in section 3.2. We get

$$\frac{\partial V(Q)}{\partial n} = \sum_{i \in L} \sum_j V_{ij} \frac{\partial S_{ij}(Q)}{\partial n} \quad (3.106)$$

and

$$V(P) = \frac{1}{4\pi} \int_{\Gamma} \left\{ \sum_{ij} V_{ij} S_{ij}(Q) \frac{\partial \frac{1}{r}}{\partial n} - \sum_{ij} V_{ij} \frac{\partial S_{ij}(Q)}{\partial n} \frac{1}{r} \right\} d\Gamma(Q) \quad (3.107)$$

The announced linear equations are thus obtained:

$$V(P) = \sum_{i \in L} \sum_j c_{ij} V_{ij} \quad (3.108)$$

In a similar manner we may even express derivatives of V at P. Differentiation of Greens formula poses no problem because P is not situated at Γ . Therefore $1(P,Q)$ never becomes zero. For example we may evaluate

$$\begin{aligned} \frac{\partial V(P)}{\partial \varphi} &= \sum_{ij} d_{ij} V_{ij} \\ \frac{\partial V(P)}{\partial \lambda} &= \sum_{ij} e_{ij} V_{ij} \\ \frac{\partial^2 V(P)}{\partial \varphi \partial \lambda} &= \sum_{ij} f_{ij} V_{ij} \end{aligned} \quad (3.109)$$

Representing $V(P)$ and the "horizontal derivatives" of V at P in this

way, and doing this for all nodes of Γ_0 , provides us with a set of linear equations which, together with the normal equations of the interior of the shell, will ensure a potential which is "reasonably harmonic" outside the earth. A disadvantage of this method, which is proposed in McDonald-Wexler (1972) and further discussed in Zienkiewicz et al. (1979), is the need for the evaluation of a number of integrals which are not all too simple. The following 3 remarks are considered important.

Remark 1: The reader may wonder why we used only V together with its horizontal derivatives in the above compatibility equations. The answer consists of two parts. First we have to point out that the underlying assumption is that of a tricubic representation of the field. The "horizontal" nodal parameters $V, V_\varphi, V_\lambda, V_{\varphi\lambda}$ are, so to speak, responsible for the variation of V along the surface Γ_0 . Secondly, we should stress that our finite set of equations substitutes for an infinite set of equations (3.107) in which P varies all over Γ_0 . The variation of P and $V(P)$ along Γ_0 is now logically replaced by the variation of those parameters in the cubic representation which are responsible for the behaviour of $V(P)$, for $P \in \Gamma_0$.

Remark 2: The discussion in Remark 1 also demonstrates an alternative way in which Greens third identity can be used in order to ensure compatibility of the field across Γ_0 . Instead of using equations (3.109) for $V_\varphi, V_\lambda, V_{\varphi\lambda}$ we could just use equation (3.108) for $V(P)$; however, we must use it 4 times as often as there are nodes on Γ_0 . A way to do this would be to use (3.108) for the nodes P on Γ_0 and also for the points halfway between two adjacent nodes. Confer figure 3.13.

x	x	x	x	x
o	x	o	x	o
x	x	x	x	x
o	x	o	x	o

o ... finite element nodes on Γ_0
 x ... additional points p at which
 Greens identity is evaluated

Figure 3.13. Pattern of nodes as
 explained in the text

Remark 3: One could even go one step further. One could use Greens identity for more horizontal parameters than this is implied by the discussion in the foregoing remarks. Exact fulfillment of the identities then can no longer be postulated. One would apply weights to the discrepancies and add their weighted sum of squares to the other functional to be minimized. In this way one even ends up with a positive definite symmetric system.

(4) Combination with spherical harmonics. Imagine the potential represented at the outer boundary Γ_0 in terms of spherical harmonics

$$V(P) = \sum_{n=0}^N \frac{1}{r_P^{n+1}} \sum_{m=-n}^{+n} C_{nm} H_{nm}(P) \quad (3.110)$$

We use the symbol "H" because the "S" is already reserved for the shape functions. We may also represent the horizontal derivatives of V in this

way

$$\begin{aligned}
 \frac{\partial V(P)}{\partial y} &= \sum_{n=0}^N \frac{1}{r_p^{n+1}} \sum_{m=-n}^{+n} C_{nm} \frac{\partial H_{nm}(P)}{\partial y} \\
 \frac{\partial V(P)}{\partial \lambda} &= \sum_{n=0}^N \frac{1}{r_p^{n+1}} \sum_{m=-n}^{+n} C_{nm} \frac{\partial H_{nm}(P)}{\partial \lambda} \\
 \frac{\partial^2 V(P)}{\partial y \partial \lambda} &= \sum_{n=0}^N \frac{1}{r_p^{n+1}} \sum_{m=-n}^{+n} C_{nm} \frac{\partial^2 H_{nm}(P)}{\partial y \partial \lambda}
 \end{aligned} \tag{3.111}$$

We form these equations for any node at Γ_0 . Imagine that there are just as many equations as there are coefficients C_{nm} . Then, hopefully, the C_{nm} could be evaluated from them. However, we do not propose to do this. Instead we propose to add equations for the remaining nodal parameters, i.e.

$$\begin{aligned}
 \frac{\partial V(P)}{\partial r} &= \sum_{n=0}^N \frac{-(n+1)}{r_p^{n+2}} \sum_{m=-n}^{+n} C_{nm} H_{nm}(P) \\
 \frac{\partial^2 V(P)}{\partial r \partial y} &= \sum_{n=0}^N \frac{-(n+1)}{r_p^{n+2}} \sum_{m=-n}^{+n} C_{nm} \frac{\partial H_{nm}(P)}{\partial y} \\
 \frac{\partial^2 V(P)}{\partial r \partial \lambda} &= \sum_{n=0}^N \frac{-(n+1)}{r_p^{n+2}} \sum_{m=-n}^{+n} C_{nm} \frac{\partial H_{nm}(P)}{\partial \lambda} \\
 \frac{\partial^3 V(P)}{\partial r \partial y \partial \lambda} &= \sum_{n=0}^N \frac{-(n+1)}{r_p^{n+2}} \sum_{m=-n}^{+n} C_{nm} \frac{\partial^2 H_{nm}(P)}{\partial y \partial \lambda}
 \end{aligned} \tag{3.112}$$

In this way, we will ensure that the potential is reasonably harmonic outside of Γ_0 compatible with the field within the shell, and tending to zero as $r \rightarrow \infty$.

Remarks 1 to 3 given above under (3) apply mutatis mutandis to the present situation.

3.8. Local data deficiencies.

It may happen that data are abundant and redundant in some areas while in others they are sparse and deficient. Lack of sufficient data in local areas can cause rank deficiencies or near singularities of the normal equations. To illustrate our point, consider Dirichlet's problem for the exterior of the unit circle:

$$\begin{aligned}\Delta V(r, \varphi) &= 0 ; & r > 1 \\ V(1, \varphi) &= f(\varphi) \\ V(r, \varphi) &= O(\log r) ; & r \rightarrow \infty\end{aligned}\tag{3.113}$$

Suppose that element partition and distribution of data are as shown in figure 3.14.

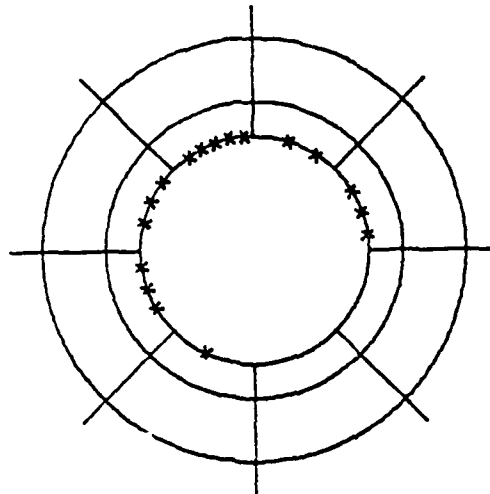


Figure 3.14. Distribution of data leading to singularities of the normal equations.

Infinite elements are used for the outer zone. It is seen that data are missing or deficient in 3 of the 8 intervals at the circumference of the unit circle. It follows that a cubic spline function $s(\varphi)$ approximating

$f(\varphi)$ in the least squares sense is undetermined. Hence the contribution of the measurements to the normal equations will have a rank smaller than in the case of sufficient data. The addition of the contribution of the field to the normal equations does not remedy the situation. The field contribution merely serves to ensure (approximate) harmonicity in the exterior region $r > 1$. If the boundary data are undetermined in Dirichlet's problem, the whole field will be undetermined.

Remark: Only to freshmen of adjustment courses it may appear paradoxical that a rank deficiency occurs when the number of observations exceeds the number of unknowns. Think of a network which may be rigid and stable in some areas and poorly determined in others.

There are at least two ways to remedy the rank deficiencies due to lack of data in some areas. (1) The elements may be chosen larger there. (2) A third contribution to the normals may be calculated which results from the square of a certain curvature norm applied to the spline $s(\varphi)$ approximating $f(\varphi)$.

Approach (1) is not too strongly recommended. It may cause a loss of regularity in the element partition. As a consequence the computer programs could become more involved.

Approach (2) is related to least squares collocation using splines. We take the finite element representation of $V(1, \varphi)$

$$V(1, \varphi) = \sum_{i \in L} \sum_j V_{ij} S_{ij}(1, \varphi) \quad (3.114)$$

The index set L comprises the nodes situated on the circumference of the unit circle. We take the second derivative

$$\frac{d^2 V(1, \varphi)}{d\varphi^2} = \sum_{i \in L} \sum_j V_{ij} \frac{d^2}{d\varphi^2} S_{ij}(1, \varphi) \quad (3.115)$$

and consider the integral

$$\frac{1}{2} \bar{E} = \int_0^{2\pi} \left[\sum_{i \in L} \sum_j V_{ij} \frac{d^2}{d\varphi^2} S(1, \varphi) \right]^2 p(\varphi) d\varphi \quad (3.116)$$

Here $p(\varphi)$ is a weight function which can be taken $p(\varphi) \equiv 1$. It may also be taken positive in areas of poor data and zero in areas of sufficient data. Taking the variation of (3.116) with respect to V_{ij} yields, very much in the same way as outlined in section 3.3.2., a third contribution to the normals. Sparsity of the normals will not be impaired in any way.

The method readily generalizes to 3 dimensions. The second derivative with respect to φ may be replaced by the surface Laplacean thereby.

4. Estimation of computation time

4.1. Nested dissection and Helmert blocking

Our normal equations can be grouped according to nodes. If d denotes the dimensions, i.e. $d = 1, 2$ or 3 and if d -cubics are used, then for each node there are 2^d rows and 2^d columns corresponding to the 2^d parameters per node. During any step in reordering or reducing the normal equations the 2^d equations for one node will always be lumped together. We may view the system as composed of $2^d \times 2^d$ submatrices which could be called generalized coefficients. Elimination is then carried out using these generalized coefficients instead of conventional scalars.

In the original normals, two nodes are coupled by nonzero off-diagonal coefficients if there is a finite element on whose boundary both nodes are located. Hence the system is sparse. During elimination the coupling increases due to fill-in. If a node is eliminated all its neighboring nodes which are not yet eliminated become coupled. The whole idea of sparse elimination is to renumber the nodes in a way that fill-in is effectively kept down. We shall use a technique which is a combination of nested dissection and Helmert blocking. The philosophy behind this procedure was discussed in great detail in Meissl (1980), sections 3.5.4 -5. Here we shall outline it on hand of a two dimensional problem with element partition shown in figure 4.1.

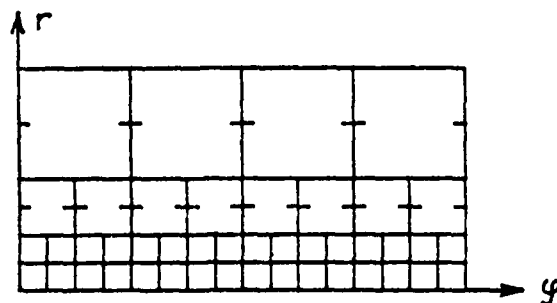


Figure 4.1. Element partition serving the outline of nested dissection in conjunction with Helmert blocking

Mind that we work in the plane polar coordinate system. Hence the region of figure 4.1 is actually a circular ring, the left and right boundary line, and the nodes located on them are to be identified. We first describe the block design in a bottom up fashion. At the first stage, we consider 8 blocks. Each one is formed out of 2 adjacent squares located at the bottom level. Such a block looks as shown in an enlarged way in figure 4.2. We imagine the normals formed for the two simple quads composing the block, and the normals added. We eliminate from them the node indicated by a circle in figure 4.2.

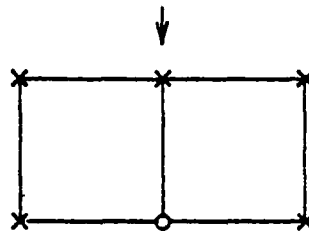


Fig. 4.2. First stage block

We call this the inner node of this block. The other nodes (indicated by crosses) are called junction nodes. We obtain a partially reduced system of 5 junction nodes.

At the second stage we enlarge the blocks of stage one by adding two adjacent squares of the next higher layer. The block looks as shown in figure 4.3. There are still 8 blocks.

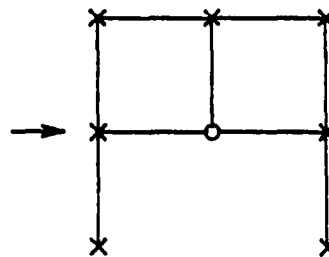


Figure 4.3. Second stage block

Each block again shows one inner node. However, this time there are 7 junction nodes. The normal equations for such a node are obtained by adding the partially reduced normals of stage 1 and the normals for the upper squares. After elimination of the inner node, we obtain a partially reduced system for the 7 junction nodes.

At stage 3 we adjoin to the block of the previous stage one compound element of the third layer. We obtain 8 blocks as shown in figure 4.4.

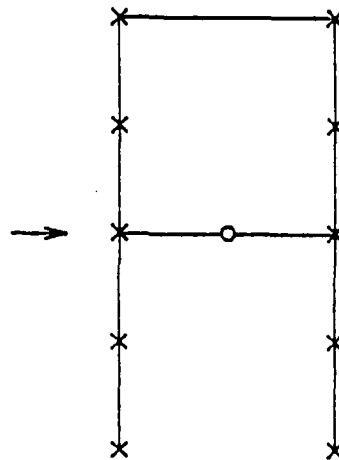


Figure 4.4. Third stage block

The normals are formed by adding the partially reduced normals of stage 2 and the normals for the compound upper quad. The inner node is eliminated.

At stage 4 we join two adjacent blocks of the previous stage and eliminate the 4 inner nodes. Confer figure 4.5. The number of blocks is 4 in this stage.

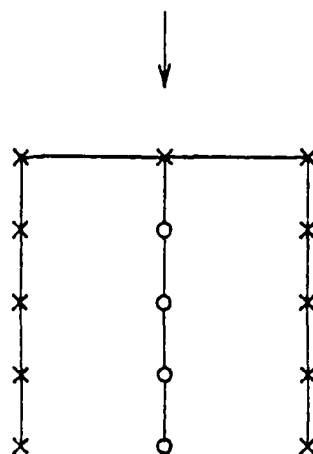


Figure 4.5. Fourth stage block

At stage 5 we adjoin to the previous block the compound elements of the upper layer. We obtain blocks as shown in figure 4.6. The scale of this figure is now the same as that of figure 4.1. The number of blocks is still 4.

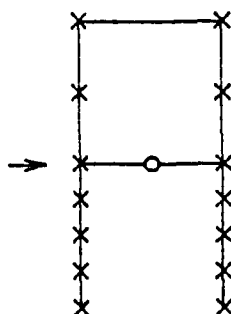


Figure 4.6. Fifth stage block

At stage 6 we lump two adjacent blocks of stage 5. See figure 4.7.

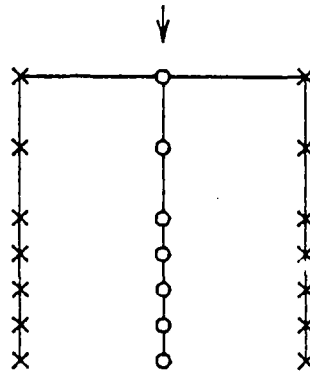


Figure 4.7. Sixth stage block

At stage 7 we do a similar thing; confer figure 4.8

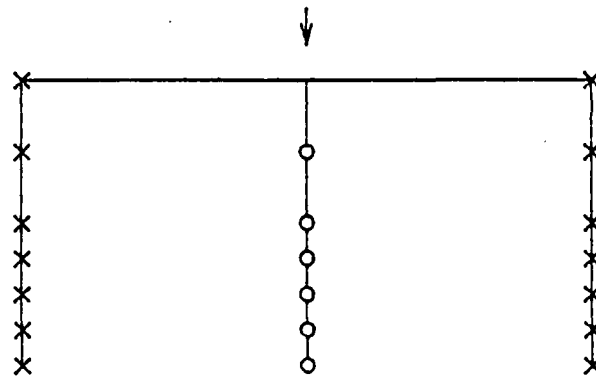


Figure 4.8. Seventh stage block

Recall that the nodes of the two outer boundaries are actually identical. In the final and 8-th stage we eliminate just these nodes. See figure 4.9

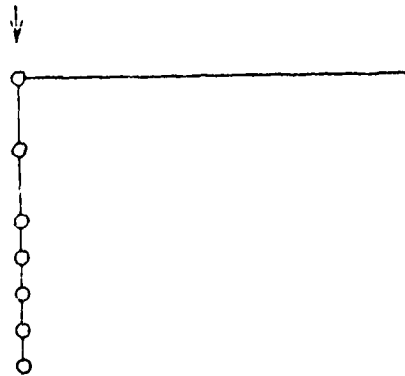


Figure 4.9. Eights stage block

The Helmert blocking procedure is, as we feel, best understood, if it is described in the above bottom up fashion. In order to analyse various block designs, however, in particular block design for 3-dimensional problems, we prefer a top down fashion, which we briefly outline for the above example.

At stage 8 we slice the spherical ring by a dissecting line. Inner nodes are those located at this line. The dissecting line is indicated by an arrow in figure 4.9.

At stage 7 we dissect the block again (arrow in figure 4.8) and we repeat this at stage 6. Inner nodes are still those nodes located at the dissecting line. At stage 5 the dissecting line is horizontal (figure 4.6). Here we discover the rule which will be general enough for all our element partitions:

Rule for identifying inner nodes and junction nodes: Inner nodes are those nodes located at the dissecting line which have not been inner nodes before. (In 3 dimensions we deal with dissecting surfaces rather than lines). Junction nodes are inner nodes of the next higher stage together with junction nodes of the next higher stage located in the block under consideration.

The dissecting lines of the stages 1-4 are easily recognized from figures 4.2 through 4.5 (arrows).

4.2. Global solution.

We consider a finite element partition which allows to model the details of the field near the surface with about the same resolution as a conventional surface layer or collocation solution based on $1^\circ \times 1^\circ$ block averages of gravity anomalies.

The dimension is 3 and we use tricubics. This implies that there are 4 parameters available at each node which are responsible for the horizontal variation of V in the vicinity of this node. Hence the element size near the surface must be chosen as $2^\circ \times 2^\circ$. Mind that the number of blocks is the same as the number of nodes. Consequently there are 4 parameters available to model the horizontal variation of V in a block. It follows that the averages over gravity anomalies in the four $1^\circ \times 1^\circ$ subblocks can be matched exactly. Our surface of computation will be a sphere. It may be imagined as a sphere slightly below the earth's surface. The computational effort of solving the normals is unaffected by the use of a more complicated reference surface. Only the formation of the normals takes longer. Asymptotically, i.e. for very large systems, formation is negligible in comparison to solution. By assuming a concentric sphere with an appropriately larger radius we obtain a spherical shell. Our element partition refers to this shell. It is specified in figure 4.10(a),(b) and in tables 4.1 to 4.3.

zone from to (degr. latit.)		block size latit. longit. (degrees)		Number of blocks along a parallel	Number of blocks in zone
0	72	2	2	180	6480
72	84	2	6	60	360
84	90	6	18	20	60

Table 4.1 Element partition in first (bottom)
layer as well as in second layer

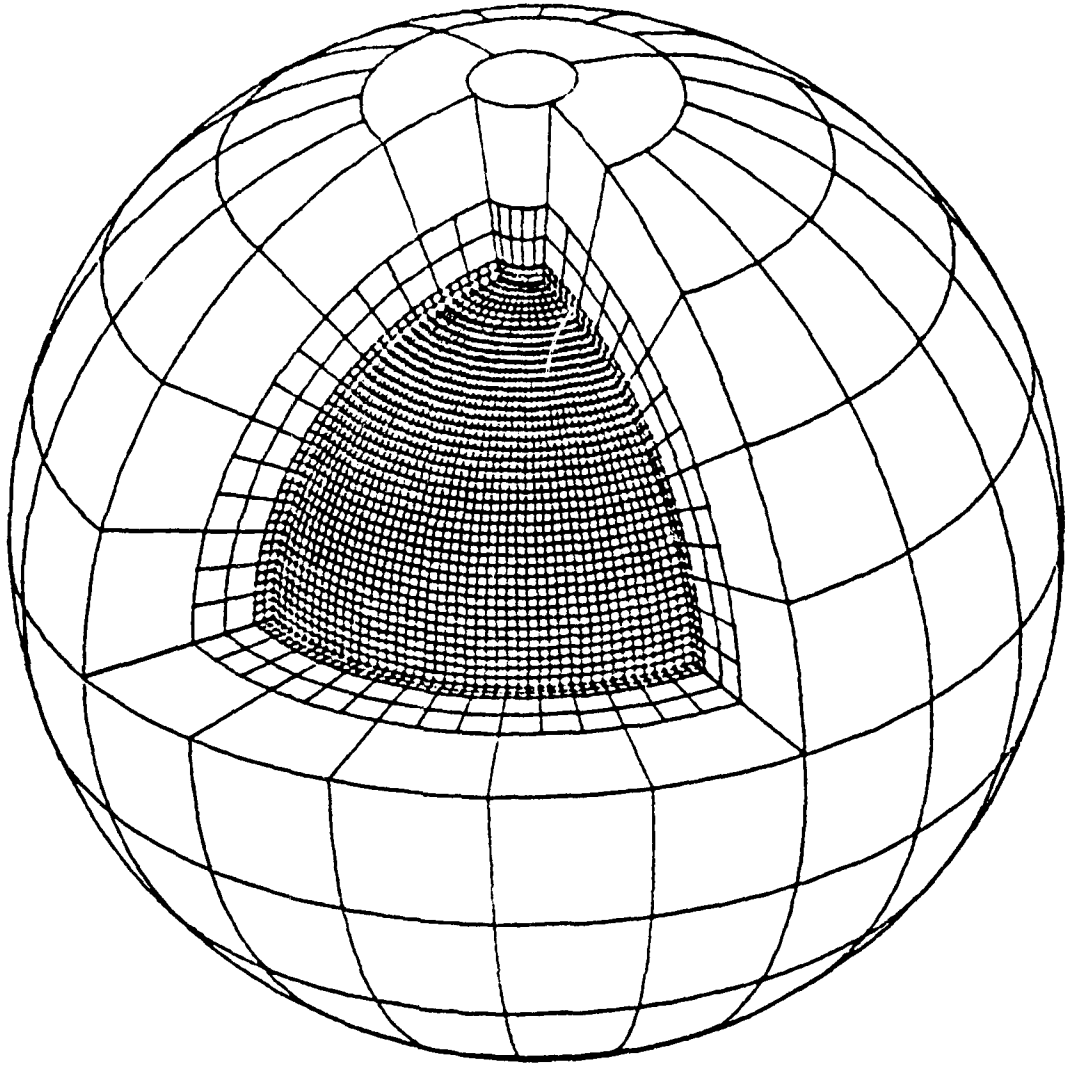


Figure 4.10(a)
Element partition for global solution

AD-A104 164

OHIO STATE UNIV COLUMBUS DEPT OF GEODETIC SCIENCE

F/G 8/5

THE USE OF FINITE ELEMENTS IN PHYSICAL GEODESY (U)

APR 81 P MEISL

F19628-79-C-0075

UNCLASSIFIED

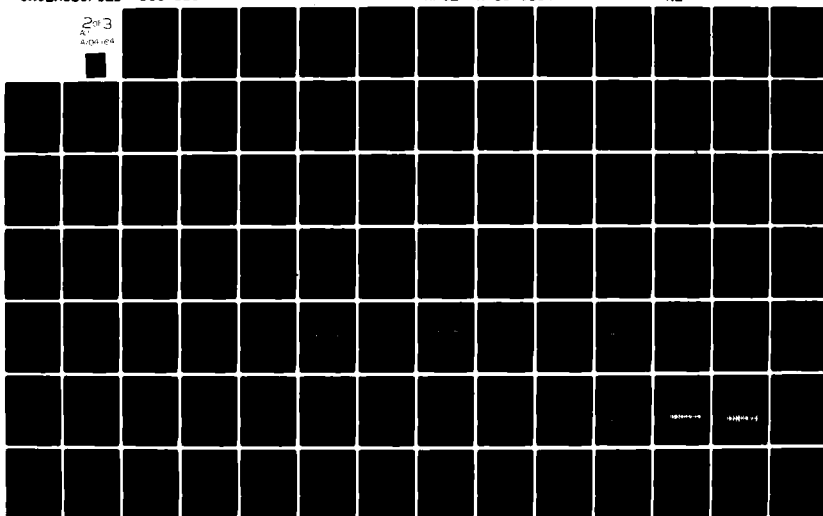
D65-313

AFGL-TR-81-0114

NL

2 of 3

A104-164



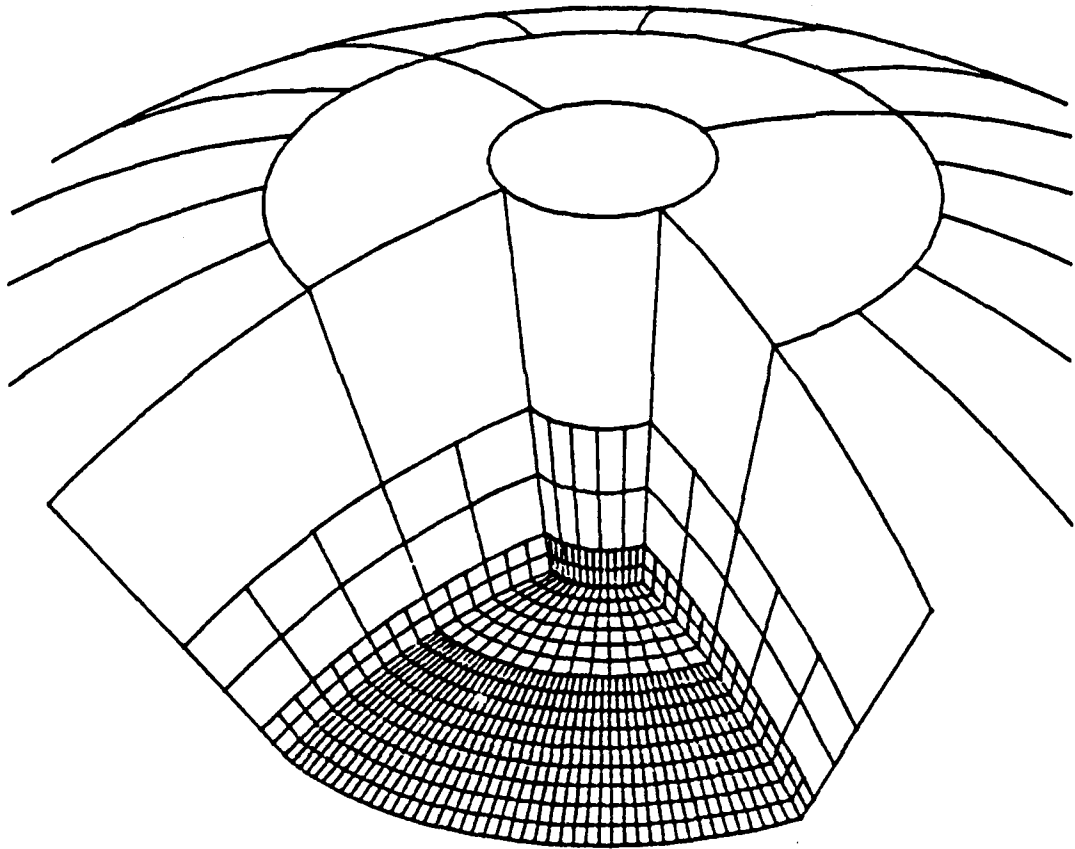


Figure 4.10(b). Element partition for global solution.
Detailed view of area near the pole

The tables apply to the northern hemisphere. The partition for the southern hemisphere follows by symmetry. In the bottom layer there are 13800 blocks over the whole sphere. Note that the area of the sphere divided by the area of a $2^\circ \times 2^\circ$ equatorial block gives about 10300.

Let us pause for a moment and reflect on the computational effort associated with a surface layer solution based on $2^\circ \times 2^\circ$ blocks. (Recall that only $1^\circ \times 1^\circ$ blocks would be about equivalent to our finite element treatment). We are dealing with a system of 13800 unknowns. The unknowns are the densities for the blocks. To solve a linear symmetric system of i unknowns by a direct elimination method requires about

$$\frac{l^3}{6} \quad (4.1)$$

elementary steps. (Confer appendix A, equation (A.6)). One elementary step comprises a multiplication followed by an addition. We arrive at

$$\frac{13800^3}{6} = 4.38 \text{ E}11 \quad (4.2)$$

steps. Assuming a computer that can perform 500,000 steps per second (OSU has such a computer at the present time), we estimate a total effort of

$$243 \text{ hours CPU time} \quad (4.3)$$

If the blocks are chosen $1^\circ \times 1^\circ$ the number of blocks multiplies by 4. Hence the estimates of (4.3) multiplies by $4^3 = 64$ giving about

$$15,000 \text{ hours CPU time} \quad (4.4)$$

We return to the finite element partition. The experience gained from the computer experiments documented in chapter 6 and dealing with the two dimensional Stokes problem, persuades us to assume two layers of elements of the size shown in table 4.1. At the third layer blocks are fused, mostly 9 into one. Table 4.2 shows the block size for the third layer. (Cf. again figure 4.10(a),(b))

zone		block size		No. of blocks	No. of blocks
from	to	lat.	long.	along a	in zone
(degr. latit.)		(degrees)		parallel	
0	72	6	6	60	720
72	84	6	18	20	40
84	90	6	90	4	4

Table 4.2. Element partition in third layer

At the fourth layer the same partition is used. At the fifth layer we fuse again; table 4.3 shows how.

zone		block size		No. of blocks	No. of blocks
from	to	lat.	long.	along a	in zone
(degr. latit.)		(degrees)		parallel	
0	72	18	18	20	80
72	84	12	90	4	4
84	90	6	360	1	1

Table 4.3. Element partition in fourth (upper) layer

Within the main zone -72° to $+72^{\circ}$ the basic building block is a $18^{\circ} \times 18^{\circ}$ configuration. We call it a 6th-stage block. The meaning of this name will become clear later on. The face of such a 6th-stage block shows a pattern of nodes as depicted in figure 4.11. We count 46 nodes on this face, and 8 along a side line.

In the following, we assume a blocking strategy for the formation and solution of the normals as outlined in the previous section. The normals result from the contribution of the field and of the observations. Confer section 3.3. It is important that the observations are local. Any observation must only involve points situated in one and the same element. The contribution of the field in the outer zone is assumed to be taken into account by finite elements of infinite size. Confer section 3.6. The partition is that one shown for the last inner layer in table 4.3.

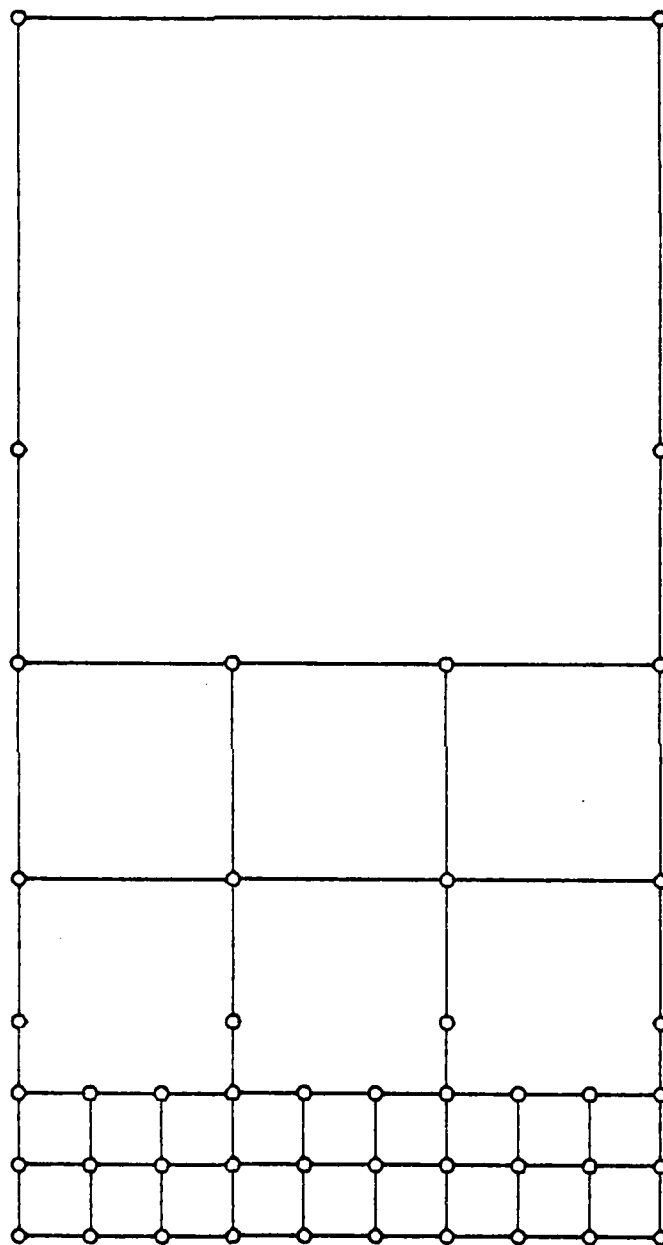


Figure 4.11. Pattern of node at a lateral
face of a stage 6-block

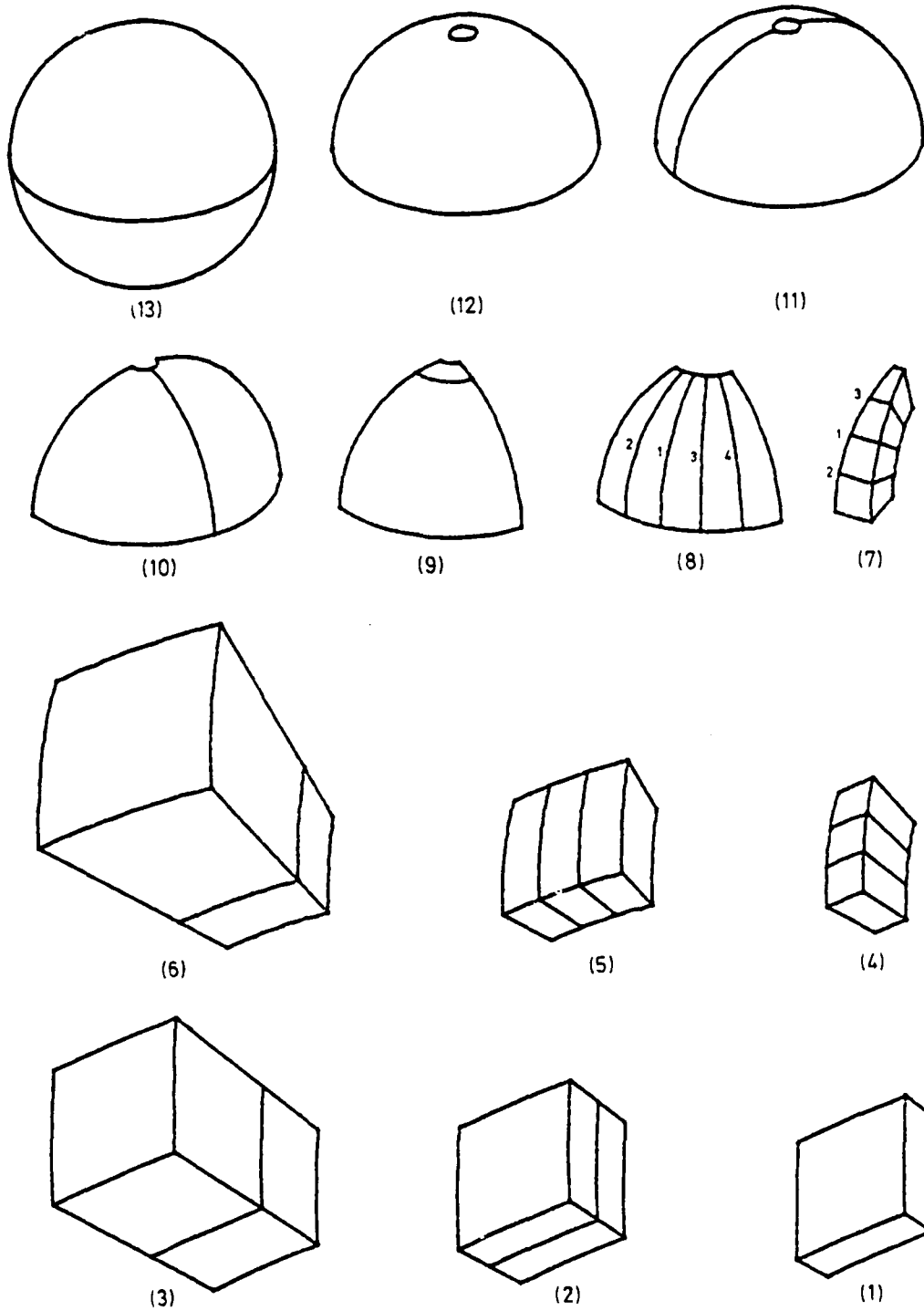


Figure 4.12. Blocking scheme for global solution (incomplete)

The blocking will be hierarchical. Block boundaries will always be composed of element boundaries. We proceed to describe the design and to estimate the computational effort in a top down manner, starting at the last stage.

At the last stage, which is stage 13, there are only inner nodes. We assume that they are comprised of all nodes situated in the equatorial plane. See figure 4.12(13). Figure 4.11. shows the pattern of nodes which repeats itself 20 times along the equator. Hence the number of nodes is

$$i = 20 \times 46 - 20 \times 8 = 760$$

The corresponding system of partially reduced normal equations has $8 \times i$ equations because we have 8 parameters per node. In agreement with equation (4.1) the computational effort in steps is

$$\frac{(8i)^3}{6} = 3.75 E10$$

translated into CPU time (assuming 500,000 steps/sec) this gives

$$\text{CPU time} = 20.8 \text{ hours} \quad [\text{stage 13}] \quad (4.5)$$

At the next lower stage, i.e. stage 12, we have 2 blocks, the northern and the southern hemisphere. We cut off the polar caps at latitudes ± 84 degrees (cf. figure 4.12(12)). We see that for one block

$$i = 248$$

whereas

$$j = 760$$

from the previous stage. According to equation (A 5) of appendix A the computational effort to eliminate n interior equations in the presence

of m junction equations is

$$\frac{n^3}{6} + \frac{nm(n+m)}{2} \quad (4.6)$$

In our case $n = 248 \times 8$, $m = 760 \times 8$. Hence the effort in steps is

$$\left[\frac{i^3}{6} + \frac{ij(i+j)}{2} \right] 8^3 \quad (4.7)$$

Translated into computer time, and multiplied by the number of blocks (=2), this gives

$$\text{CPU time} = 55.5 \text{ hours} \quad [\text{stage 12}] \quad (4.8)$$

We proceed to the next lower stage, i.e. stage 11. We are still dealing with 2 blocks. Each block is divided into two parts by the Greenwich meridian. Keeping figure 4.12(11) in mind we count in one block

$$i = 8 \times 46 + 2 \times 34 - 12 \times 8 = 340$$

interior nodes (along the dividing meridian) and

$$j = 760 + 248 = 1008$$

junction nodes (at $\varphi = 0^\circ, \pm 84^\circ$). The CPU for both blocks is estimated at

$$\text{CPU time} = 135.1 \text{ hours} \quad [\text{stage 11}] \quad (4.9)$$

At the next lower stage, which is stage 10, we are dealing with 4 identical blocks. Cf. figure 4.12(10). Each one comprises a quarter of the sphere. We subdivide each block by the central meridian. For one such block we have

$$i = 4 \times 46 + 34 - 6 \times 8 = 170$$

$$j = 10 \times 46 + 2 \times (4 \times 46 + 34) + 2 \times 70 - 22 \times 8 = 860$$

We obtain a total CPU for this stage as

$$\text{CPU time} = 86.6 \text{ hours} \quad [\text{stage } 10] \quad (4.10)$$

At the 9-th stage we have 8 blocks corresponding to the 8 octants. We cut off the region near the pole by dissecting along the surface $\varphi = \pm 72^\circ$. See figure 4.12(9). We have

$$i = 5 \cdot 46 - 6 \cdot 8 = 182$$

$$j = 13 \cdot 46 + 2 \cdot 34 + 70 - 16 \cdot 8 = 608$$

we obtain

$$\text{CPU time} = 101.8 \text{ hours} \quad [\text{stage } 9] \quad (4.10a)$$

From now on we describe the details only for the area bounded by the latitudes $\pm 72^\circ$. The portions enclosed by latitudes $72^\circ < \varphi < 84^\circ$, $-72^\circ > \varphi > -84^\circ$ will be treated summarily later on.

At stage 8 we cut the octant (truncated at $\varphi = \pm 72^\circ$) into 5 slices as shown in figure 4.12(8). The meridians are eliminated in a sequence implied by the figure. The effort results from the following table 4.4

No.	i	j	CPU
=====	=====	=====	=====
1	144	684	93.9
2	144	456	45.9
3	144	532	60.1
4	144	456	45.9

Table 4.4. Computational effort
for the 4 phases of stage 8

Summing for this stage we obtain

$$\text{CPU time} = 245.8 \text{ hours} \quad [\text{stage } 8] \quad (4.11)$$

At stage 7 we have 40 blocks in the shape of slices as shown in figure 4.12(7). Each slice is split into 4 blocks. The elimination of the dividing lines proceeds in a sequence as shown in figure 4.12(7). The effort results from the following table 4.5

No.	i	j	CPU
=====			
1	30	380	26.6
2	30	228	10.1
3	30	228	10.1

Table 4.5. Computational effort
for the 3 phases of stage 7

We obtain for this stage a

$$\text{CPU time} = 46.8 \text{ hours} \quad [\text{stage 7}] \quad (4.12)$$

At stage 6 we are dealing with 160 $18^\circ \times 18^\circ$ blocks. We split one block into 2 parts by a spherical surface between the 3rd and 4th layer. Confer figure 4.12(6). We have for one block

$$i = 4$$

$$j = 4 \times 46 - 4 \times 8 = 152$$

The resulting CPU for all 160 blocks is

$$\text{CPU time} = 2.2 \text{ hours} \quad [\text{stage 6}] \quad (4.13)$$

We ignore the upper blocks in the sequel. Their contribution is marginal.

At stage 5 the number of essential blocks is still 160. We decompose anyone of them into 3 subblocks of size 6 degrees by 18 degrees. Cf. figure 4.12(5). There are two phases, and each of them has

$$i = 26$$

whereas

$$j = 4*38 - 4*5 + 2 = 134$$

$$j = 2*38 + 2*27 - 4*5 = 110$$

The effort is

$$\text{CPU time} = 21.8 \text{ hours} \quad [\text{stage 5}] \quad (4.14)$$

At stage 4 we have 480 blocks. (Figure 4.12(4)). Each one is subdivided into three $6^\circ \times 6^\circ$ blocks. Again there are 2 phases:

$$i = 6$$

$$j = 2*27 + 2*16 - 4*5 = 66$$

$$j = 2*38 + 2*16 - 4*5 = 88$$

$$\text{CPU time} = 5.4 \text{ hours} \quad [\text{stage 4}] \quad (4.15)$$

At stage 3 the number of blocks is 1440. (Figure 4.12(3)). We eliminate the 4 nodes at the lower face of the upper subblock. Consequently we have for one block:

$$i = 4$$

$$j = 4*16 - 4*5 = 44$$

$$\text{CPU time} = 1.7 \text{ hours} \quad [\text{stage 3}] \quad (4.16)$$

The upper portion cut off in this way does not have any interior nodes and is therefore ignored.

At stage 2 (see figure 4.12(2)) we eliminate the 4 nodes situated below those eliminated in the previous stage:

$$i = 4$$

$$j = 4*12 - 4*3 + 4 = 40$$

$$\text{CPU time} = 1.5 \text{ hours} \quad [\text{stage 2}] \quad (4.17)$$

At the first stage the 4 inner nodes at the bottom of each of the 1440 blocks are eliminated (see figure 4.12(1)).

$$i = 4$$

$$j = 4 \times 8 - 4 \times 2 + 4 = 28$$

$$\text{CPU time} = 0.7 \text{ hours} \quad [\text{stage 1}] \quad (4.18)$$

We now return to the chips of the octants cut off in stage 9. They are located between ± 72 and ± 84 degrees of latitudes. Their structure is best seen from figure 4.10(b). Table 4.6 summarizes the elimination stages necessary to decompose these blocks.

Stage	Dissecting surface	i	j	No. blocks	CPU
8'	spherical between 3-rd and 4-th layer ¹⁾	4	304	8	0.4
7'	$\lambda = \text{const.}$, producing 5 slices	16	264	8	1.4
		16	132	8	0.4
		16	176	8	0.6
		16	132	8	0.4
6'	$y = \pm 78^\circ$	6	81	40	0.2
5'	as in stage 3 ²⁾	4	44	80	0.1
4'	as in stage 2	4	40	80	0.1
3'	as in stage 1	4	28	80	0.0
$\Sigma = 3.6$					

1) Upper portion ignored in the sequel

2) blocks of stage 5' and 3 are identical. They are $6^\circ \times 6^\circ$ blocks involving layers 1,2,3.

Table 4.6. Complementary stages for regions at high latitudes

Hence we obtain

CPU time = 3.6 hours [stages 8' to 3']

The total CPU is obtained by summing over all stages.

Total CPU time = 729 hours (4.19)

Remark: We have been somewhat wasteful by allowing many elements of small width (6 degrees longitude) at latitudes $\varphi = \pm 84$. One could at the lower layers imagine an intermediate ring of elements between 84 and 86 degrees latitudes whereby the longitudinal width would be 18 degrees. This would probably reduce the total CPU time to about 600 hours.

In any case it turns out that the problem is not managable on a computer doing only half a million steps per second. However there are faster computers. The ILLIAC IV is described in Avila et.al. (1978) as a machine that has 64 processors, each of a speed comparable to that one assumed above. All processors execute the same instruction at a time, but each one operates on a separate data stream. Data can be exchanged between processors (a subset of the processors can be disabled under program control). Such a machine appears to be well suited to cut down the CPU time by a factor which may approach 64. The elimination procedure outlined above can be viewed as a procedure eliminating nodes, whereby each node contributes 8 parameters to the system. We may thus view our linear system as one having as many unknowns as there are nodes, however each unknown represents actually a subvector of 8 elements. From the viewpoint of eliminating nodes, an elementary operation is thus a multiplication of two 8×8 matrices followed by an addition. There should be a way to organize these operations such that all 64 processors of the ILLIAC IV are busy all the time.

It appears, after all, that a global solution is feasible if one of the worlds best computers is available. However, I consider it doubtful that such a computational adventure will be undertaken in the next future. Fortunately enough there is another bonus available in physical geodesy, as well as in other disciplines, which allows us to calculate a detailed and good solution in a local area of interest

without being forced to calculate such a detailed solution everywhere. The effect which we are going to exploit, and which has been exploited a lot in the past, is called "the remote zone effect". It is called, by the way, "St. Venant's" principle in elasticity theory.

4.3. The remote zone effect.

There is not much need to elaborate on it at length. Everybody knows that the details of the field in one area have little correlation with the details in remote areas. Put it in other words, if the field is changed in an area e.g. by changing the mass distribution in this area, the high frequent features in remote areas remain practically unchanged. In applications of Stokes and Vening Meinesz formulas this effect is utilized in a way, that detailed data are only processed in a fairly small area around the point of interest. The remote zone effect can also be exploited in conjunction with other methods, e.g. the surface layer method, the collocation method and, of course, the finite element method. We shall consider two configurations. The first one is a strip in which a detailed geoid is sought. The second one is a local region, which for simplicity will be taken as rectangularly shaped.

4.4. Detailed solution in a strip.

Consider a partition at ground level as shown in figure 4.13(a). Assume that this pattern extends around the globe, and that the central line corresponds to the equator. The size of the smallest squares is assumed to be $2^\circ \times 2^\circ$. The partitioning of the space outside the surface (assumed spherical) is demonstrated in fig. 4.13(b). The size of the elements increases as we go away from the surface. The outer zone is assumed to be partitioned into specially designed elements which share two more nodes at the north- and south-pole. See also fig. 4.14.

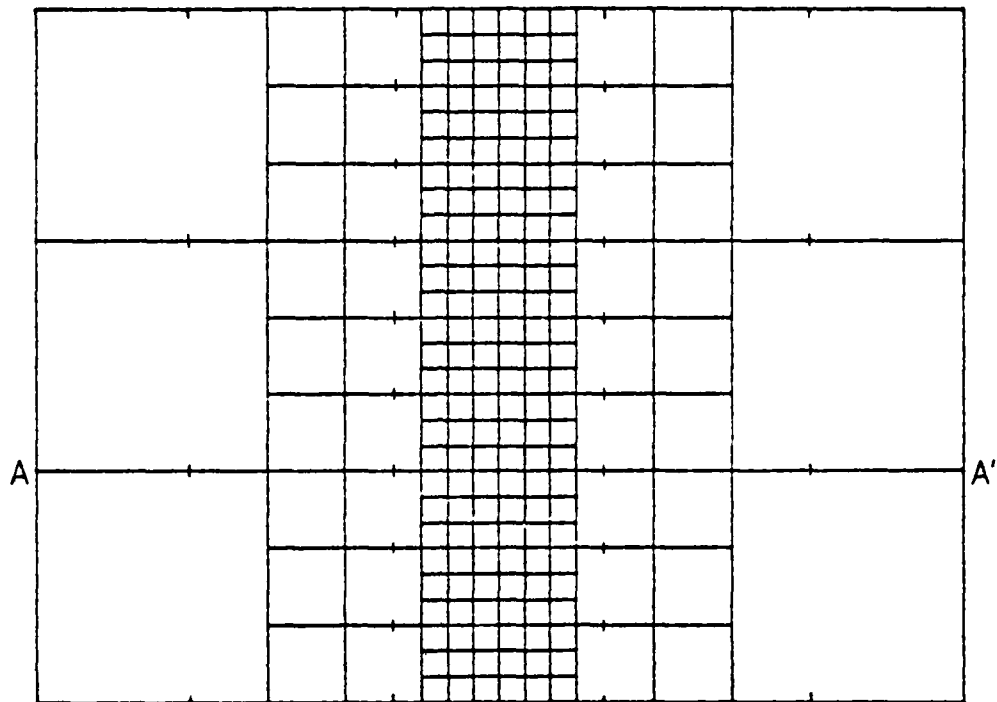


Figure 4.13(a). Partition at ground level of strip

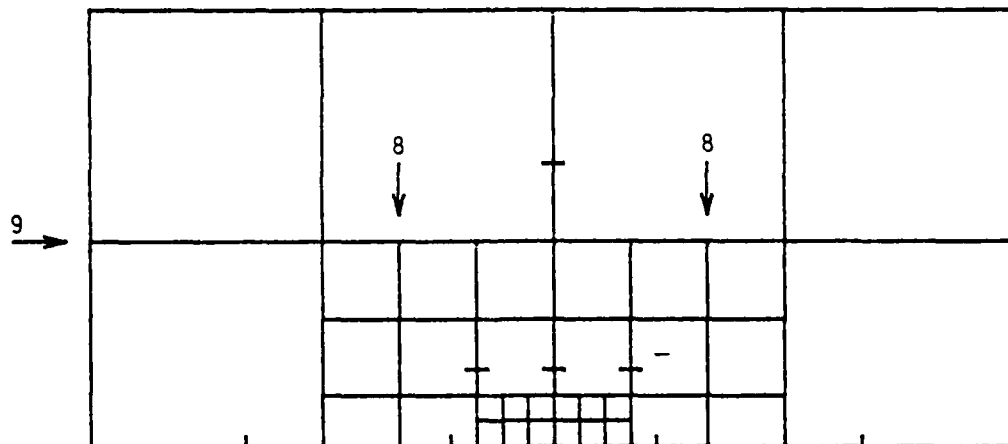


Figure 4.13(b). Partition of vertical cross section AA' as indicated in figure 4.12(a).

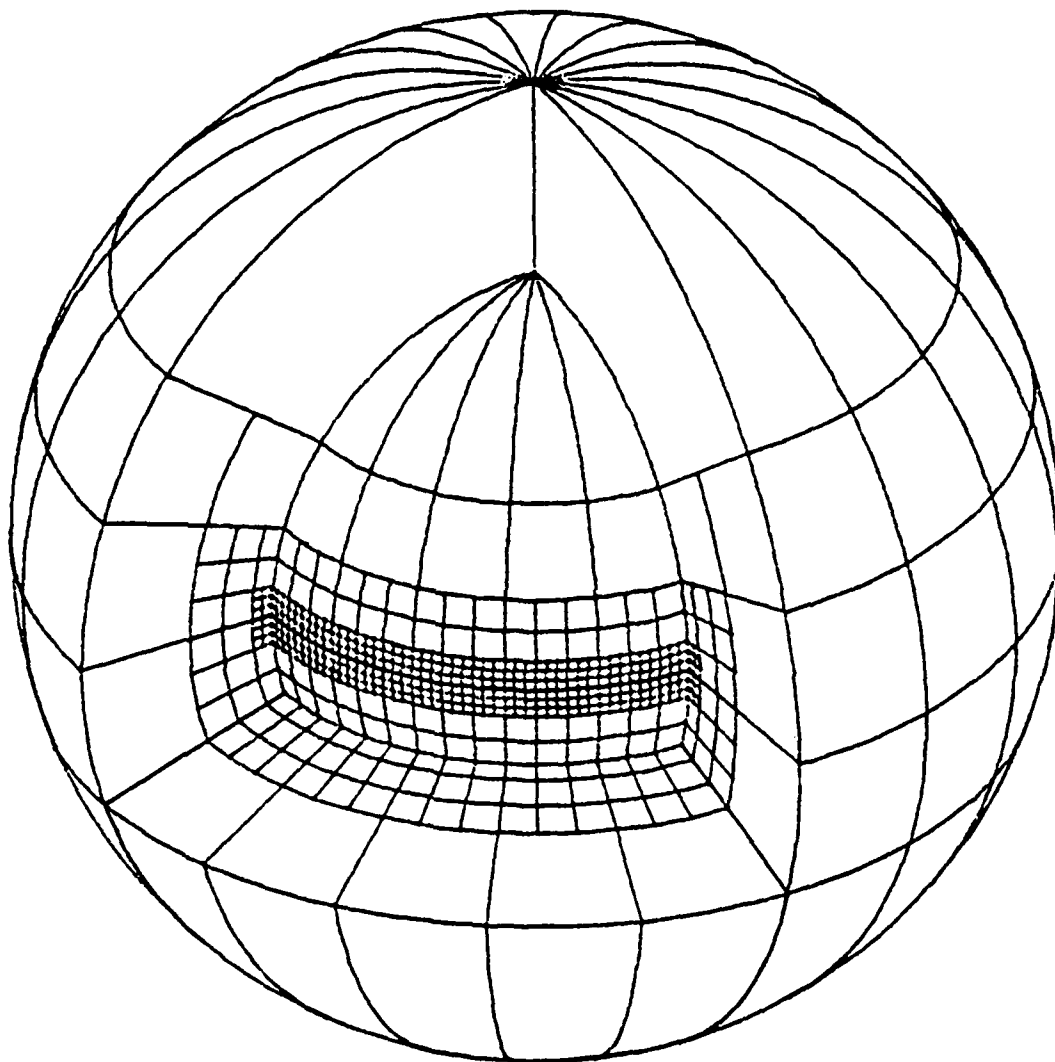


Figure 4.14. Element partition for solution
in equatorial strip

The potential at the two nodes at north- and south-pole will be assumed to be known. Only a very much smoothed version of the potential is needed there. Hence these nodes will not contribute to the

computational effort.

The idea is to get a detailed field in the vicinity of the nodes located at a vertical plane passing through the axis of the strip.

At the final stage (stage 10), we eliminate all nodes located at the main profiles. There are 20 such profiles, each having 60 nodes. We obtain a system which is structured as shown in figure 4.15

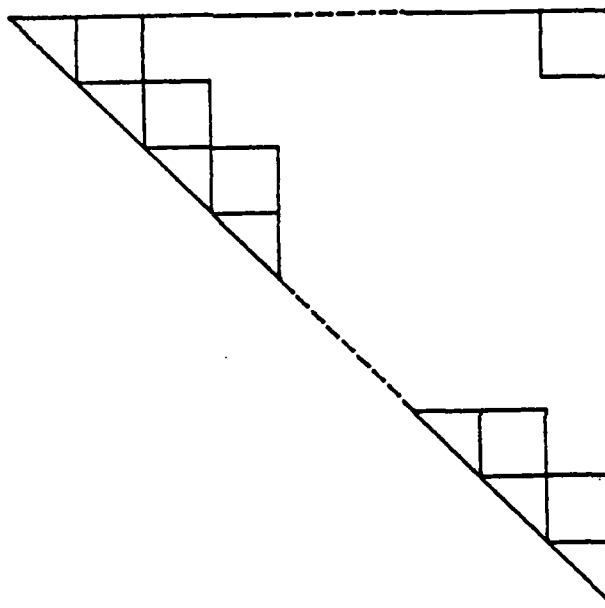


Figure 4.15. Cyclic block banded structure of normal equations at the last stage of the strip solution.

We are dealing with a cyclically blockbanded system. There are $n = 20$ diagonal blocks. The size of each block, denoted m , results from the 60 nodes at one profile, i.e. $m = 60 \times 8 = 480$. The effect to triangularize such a system amounts approximately to

$$\frac{8}{3} nm^3 \quad (4.20)$$

steps. Under our usual assumptions the resulting CPU time is

$$\text{CPU time} = 3.28 \text{ hours} \quad [\text{stage } 10] \quad (4.21)$$

At the next lower stage, which is stage 9, we are dealing with 20 blocks. We cut each into two parts by a spherical surface indicated in figure 4.13(b) by a horizontal arrow labeled 9. We have $n = 20$ blocks, each one having $i = 14$, $j = 2 \times 60 = 120$. The resulting effort amounts to

$$n \left[\frac{i^3}{6} + \frac{ij(i+j)}{2} \right] 8^3 \quad (4.22)$$

steps. The

$$\text{CPU time} = 0.64 \text{ hours} \quad [\text{stage } 9] \quad (4.23)$$

No further effort is required to deal with the upper part cut off at stage 9. There are no interior nodes left in it.

At the next stage we cut off two portions at the outside of any block by considering two vertical cones (surfaces of $\varphi = \pm 12^\circ$ and indicated by vertical arrows labeled 8 in figure 4.13(b)). We have $n = 20$ (number of blocks), $i = 12$, $j = 2 \times 54 = 108$. We get

$$\text{CPU-time} = 0.44 \text{ hours} \quad [\text{stage } 8] \quad (4.24)$$

From now on (stage 7) we deal with $n = 20$ blocks, each one having a length of 18° and a profile as shown in figure 4.16

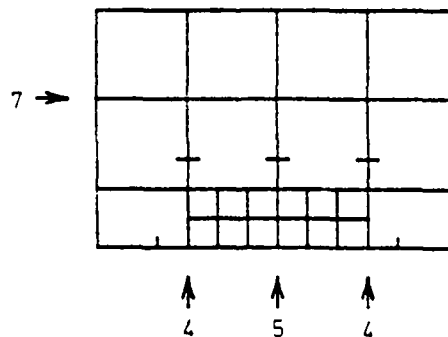


Figure 4.16. Vertical block-profile at stage 7.

We cut off the upper level by a spherical surface indicated by an arrow labeled 7 in figure 4.16. We have $i = 6$, $j = 2 \times 40 + 2 \times 11 = 102$. Hence

$$\text{CPU-time} = 0.19 \text{ hours} \quad [\text{stage 7}] \quad (4.25)$$

At stage 6 we cut each of the 20 blocks into 3 parts by 2 planes vertical to the ground level and the axis. There are two steps, each one having $i = 26$. The number of junction nodes is $j = 79$ and $j = 70$ respectively.

$$\text{CPU-time} = 1.14 \text{ hours} \quad [\text{stage 6}] \quad (4.26)$$

At stage 5 we are dealing with $n = 60$ blocks. We bisect by a vertical axial plane. See the vertical arrow marked 5 in figure 4.16. We count $i = 6$, $j = 2 \times 35 = 70$. The number of blocks is $n = 60$. Hence

$$\text{CPU time} = 0.27 \text{ hours} \quad [\text{stage 5}] \quad (4.26a)$$

At stage 4 we remove the lateral portion of each of the 120 blocks. See the arrows labeled 4 in figure 4.16. We find $i = 6$, $j = 2 \times 20 + 6 = 46$

$$\text{CPU time} = 0.25 \text{ hours} \quad [\text{stage 4}] \quad (4.27)$$

The profile of a stage 3 block looks as shown in figure 4.17. The depth of such a block is 6 degrees of longitude. Hence the horizontal cross section is square shaped. We are still dealing with 120 blocks. At each of the subsequent 3 stages we cut by a spherical surface. See the arrows labeled 3,2,1 in figure 4.17.

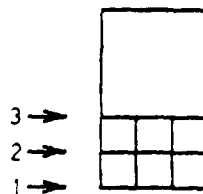


Figure 4.17. Vertical block-profile at stage 3

At stage 3 we have: $i = 4$, $j = 4 \times 14 - 4 \times 4 = 40$

$$\text{CPU-time} = 0.12 \text{ hours} \quad [\text{stage 3}] \quad (4.28)$$

At stage 2 we have: $i = 4$, $j = 4 \times 12 - 4 \times 3 = 36$

$$\text{CPU-time} = 0.10 \text{ hours} \quad [\text{stage 2}] \quad (4.29)$$

At stage 1 we have: $i = 4$, $j = 4 \times 8 - 4 \times 3 = 20$

$$\text{CPU-time} = 0.03 \text{ hours} \quad [\text{stage 1}] \quad (4.30)$$

The total CPU time is

$$\text{Total CPU-time} = 6.46 \text{ hours} \quad (4.31)$$

4.5. Rectangular region.

Consider a region of interest covering an area of $32^\circ \times 64^\circ$ (the conterminous United States are contained in such a region). Augment this region by adding layers of successively larger elements to account for the remote zone effect. The element design at ground level is seen from figure 4.18. Note that the element size at surface level is now $1^\circ \times 1^\circ$.

The vertical design is based on 4 layers of elements. The lower two layers follow the pattern of figure 4.18. At the third layer we fuse 4 blocks into one, and we do the same at the fourth layer. The basic building block of our element design is thus a block, which we call 5-stage block. The element partition at one of its four vertical faces is shown in figure 4.19 (a). There are 25 nodes on a face and 7 nodes at a vertical edge. The corresponding blocks for the augmented zones are shown in figure 4.19 (b), (c).

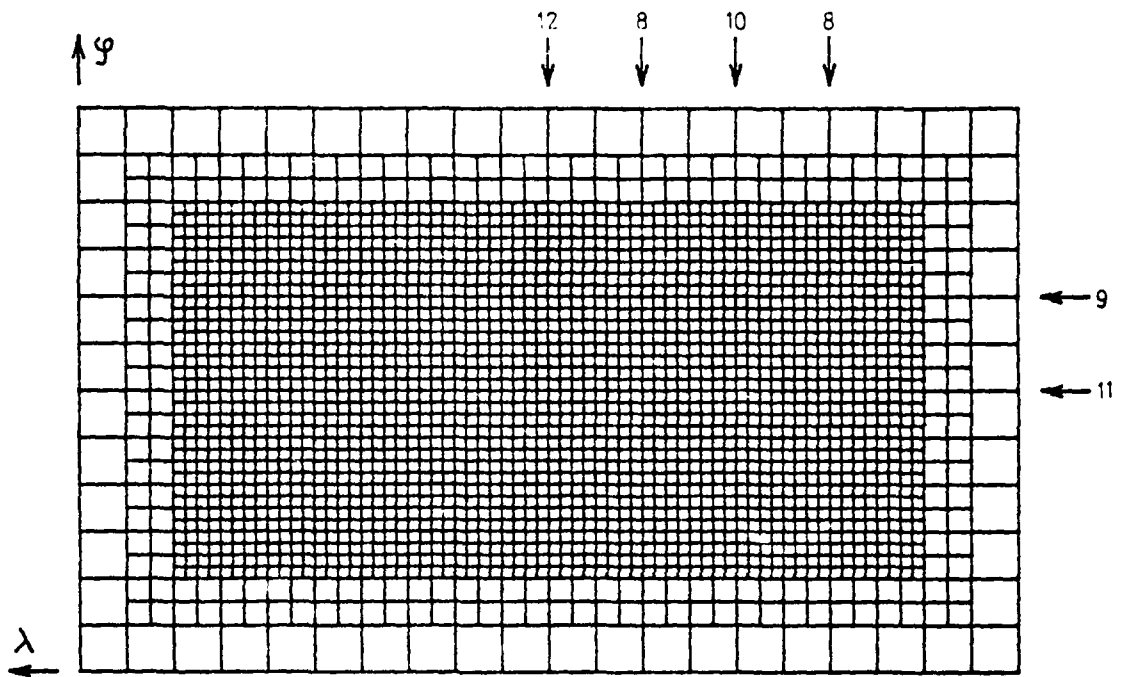


Figure 4.18. Element partition at ground level of rectangular region

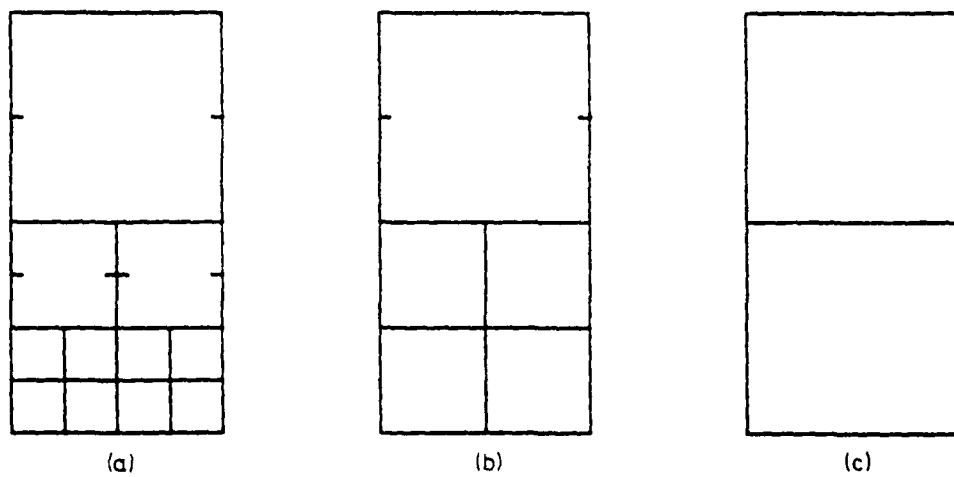


Figure 4.19. Element partition at vertical block faces in solution for rectangular region.

At the final stage (stage 12) we imagine a bisecting vertical plane passing through the central meridian. See the arrow labeled 12 in figure 4.18. We count

$$i = 8*25 - 7*7 + 2*13 - 2*5 + 2*6 - 2*3 = 173$$

Hence

$$\text{CPU time} = 0.25 \text{ hours} \quad [\text{stage 12}] \quad (4.32)$$

At stage 11 we split along the central parallel. Accordingly

$$i = 8*25 - 8*7 + 1*13 - 5 + 1*6 - 3 = 155$$

whereas

$$j = 173$$

There are two blocks

$$\text{CPU time} = 2.85 \text{ hours} \quad [\text{stage 11}] \quad (4.33)$$

At stage 10 we have 4 blocks. Any of them is split into two subblocks along its central meridian:

$$i = 4*25 - 4*7 + 1*13 - 1*5 + 1*6 - 1*3 = 83$$

$$j = 12*25 - 11*7 + 2*8 + 2*3 = 245$$

$$\text{CPU time} = 3.90 \text{ hours} \quad [\text{stage 10}] \quad (4.34)$$

At stage 9 we have 8 blocks. We split by parallels. There are 2 types of blocks. For four blocks we have

$$i = 4*25 - 5*7 = 65$$

$$j = 12*25 - 11*7 + 2*8 + 2*3 = 245$$

$$\text{CPU time} = 2.86 \text{ hours}$$

For four blocks we have

$$i = 4 \times 25 - 4 \times 7 + 1 \times 8 + 1 \times 3 = 83$$

$$j = 8 \times 25 - 7 \times 7 + 2 \times 8 + 2 \times 3 = 173$$

$$\text{CPU time} = 2.20 \text{ hours}$$

Summing for this stage we get

$$\text{CPU time} = 5.06 \text{ hours} \quad [\text{stage 9}] \quad (4.35)$$

At stage 8 we split again by meridians. See table 4.7

No. of blocks	i	j	CPU (hours)
=====			
4	29	216	0.88
4	47	173	1.04
4	29	209	0.83
4	47	137	0.69
=====			

$$\Sigma = 3.44 \text{ hours}$$

Table 4.7. Contribution of various
block types at stage 8

We see

$$\text{CPU time} = 3.44 \text{ hours} \quad [\text{stage 8}] \quad (4.36)$$

In the following stages we ignore the boundary layers. Our estimation procedure will give an upper bound on the CPU.

At stage 7 we assume 32 blocks of size $8^\circ \times 8^\circ$. Splitting in half means $i = 29$, $j = 144$

$$\text{CPU time} = 3.32 \text{ hours} \quad [\text{stage 7}] \quad (4.37)$$

At stage 6 we have 64 blocks, $i = 11$, $j = 108$

$$\text{CPU time} = 1.29 \text{ hours} \quad [\text{stage 6}] \quad (4.38)$$

At stage 5 we have 128 blocks of size $4^\circ \times 4^\circ$. We remove the upper layer:
 $i = 1, j = 72$

$$\text{CPU time} = 0.10 \text{ hours} \quad [\text{stage 5}] \quad (4.39)$$

At stage 4 we have still 128 blocks. We bisect by a vertical plane: $i = 10, j = 4 \times 21 - 4 \times 5 + 1 = 65$

$$\text{CPU time} = 0.89 \text{ hours} \quad [\text{stage 4}] \quad (4.40)$$

At stage 3 we bisect the 256 blocks once more: $i = 3, j = 2 \times 16 + 2 \times 8 = 48$

$$\text{CPU time} = 0.27 \text{ hours} \quad [\text{stage 3}] \quad (4.41)$$

Now we have 512 blocks of size $2^\circ \times 2^\circ$. We remove the upper half: $i = 1, j = 4 \times 8 = 32$

$$\text{CPU time} = 0.08 \text{ hours} \quad [\text{stage 2}] \quad (4.42)$$

At stage 1 we remove the two lowest nodes at the central axis of a block: $i = 2, j = 4 \times 9 - 4 \times 3 + 1 = 25$

$$\text{CPU time} = 0.10 \text{ hours} \quad [\text{stage 1}] \quad (4.43)$$

Summing up we obtain:

$$\text{Total CPU time} = 21.55 \text{ hours} \quad (4.44)$$

This is still quite large. On the other hand, if one owns a computer, such computation times are not unrealistic.

4.6. Further effort to cut down the computation time.

Examining the above models, in particular the global one and the rectangular one, one realizes that most of the computation time is spent

at the higher stages, not counting the final stage. The large computation time is partly due to the large number of junction nodes encountered at these critical stages.

There appears to be a possibility to further cut down on the computational effort by more sophisticated use of the remote zone effect (St. Venants principle). Let us outline the procedure on hand of our last example.

Suppose that we first attempt a strip solution, but one with an insufficient number of elements at the lateral and upper part of the strip. To be specific, suppose that we deal with the strip along the central parallel in figure 4.18 and that we use an element partition shown in figure 4.20(a) (horizontal partition at ground level) and figure 4.20(b) (vertical profile).

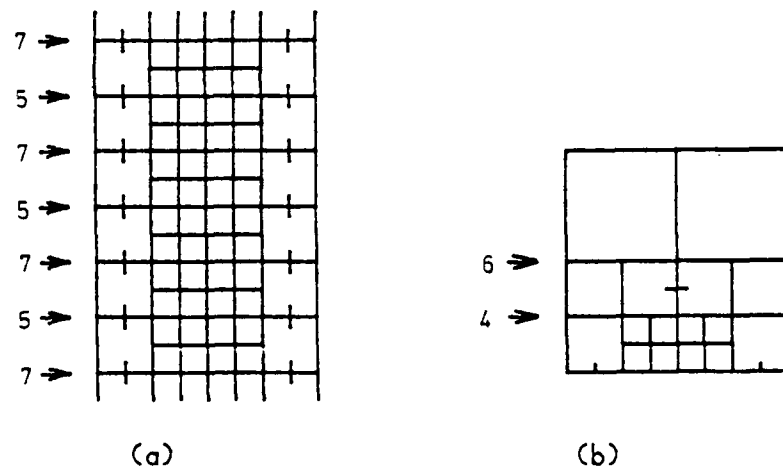


Figure 4.20
Element partition for auxiliary strip solution
(a) horizontal section, (b) vertical section

At the final stage (stage 7) we deal with a symmetric block-tridiagonal system. In contrast to figure 4.15 it is not cyclical. The effort for n blocks of size m is $\frac{7}{6}nm^3$. In our case $n = 16+1 = 17$, $m = 30 \times 8$. (We have increased n by 1 in order to approximately account for the contribution

of the boundary elements). Hence

$$\text{CPU time} = 0.15 \text{ hours} \quad [\text{stage 7}] \quad (4.45)$$

At stage 6 we cut off the upper blocks: $i = 5$, $j = 60$, number of blocks = 16

$$\text{CPU time} = 0.04 \text{ hours} \quad [\text{stage 6}] \quad (4.46)$$

At stage 5 we cut into halves by a vertical plane orthogonal to the axis: $i = 22$, $j = 2 \times 27 + 5 = 59$, number of blocks = 16

$$\text{CPU time} = 0.25 \text{ hours} \quad [\text{stage 5}] \quad (4.47)$$

At stage 4 we remove the upper blocks: $i = 5$, $j = 54$, number of blocks = 32

$$\text{CPU time} = 0.07 \text{ hours} \quad [\text{stage 4}] \quad (4.48)$$

At stage 3 we remove the lateral elements. We have $i = 4$, $j = 2 \times 21 + 5 = 47$, number of blocks = 32

$$\text{CPU time} = 0.04 \text{ hours} \quad [\text{stage 3}] \quad (4.49)$$

We have 32 blocks shaped as in figure 4.21

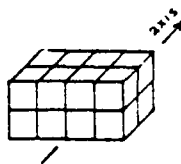


Figure 4.21. Block at stage 2

We cut into halves by a vertical plane along the axis: $i = 2$, $j = 39$, number of blocks = 32

$$\text{CPU time} = 0.01 \text{ hours} \quad [\text{stage 2}] \quad (4.50)$$

At stage 1 we remove the two remaining inner nodes: $i = 2$, $j = 25$,

number of blocks = 64

$$\text{CPU time} = 0.01 \text{ hours} \quad [\text{stage 1}] \quad (4.51)$$

Hence

$$\text{Total CPU-time} = 0.57 \text{ hours} \quad (4.52)$$

It is seen that such a profile can be calculated in about 0.5 hours. Profiles in meridional direction require about 0.25 hours.

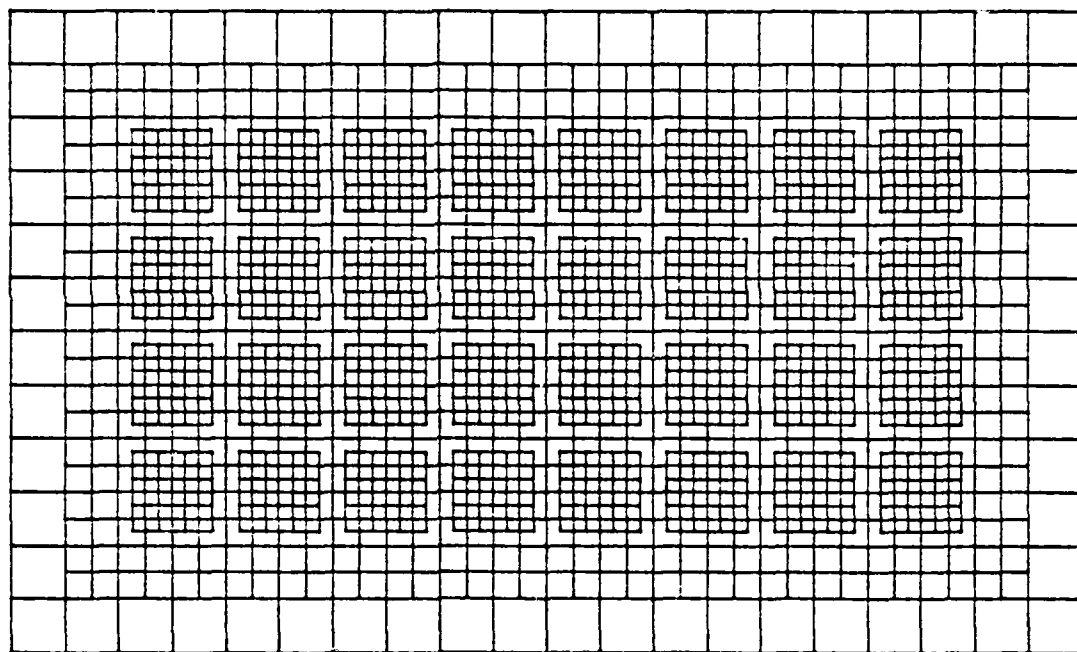


Figure 4.22. Partition of rectangular region after auxiliary profile solutions have been calculated

We calculate 3 long and 7 small profiles as shown in figure 4.22. The

$$\text{CPU time} = 3.72 \text{ hours}$$

The potential and its nodal derivatives are not thought to be known precisely at the axial plane after a profile calculation. However the

high frequent portion is thought to be known with good accuracy. This means that in the further treatment of the rectangular region we may go along with a smaller number of junction nodes at the vertical planes along profiles. The number of junction nodes along a profile follows from figure 4.19 (b). Considering this reduced number, the computation time for the various stages of section 4.5 is now reduced as shown in table 4.8.

Stage	i	j	No.	CPU
=====				
12 (final)	91		1	0.04
11	75	91	2	0.36
10	43	123	4	0.51
9	27	123	4	0.29
	43	91	4	0.31
8	11	96	4	0.06
	27	91	4	0.17
	11	107	4	0.08
	27	75	4	0.12
7	29	64	32	0.82
6	11	66	64	0.51
5 (estimated from earlier values)				< 0.10
4	10	50	128	0.55
3 (estimated from earlier values)				< 0.27
2	-	-		< 0.08
1	-	-		< 0.10
=====				
				$\Sigma < 4.37$

Table 4.8. Reduced computational effort for the solution in a rectangular region after the calculation of auxiliary profiles

We obtain a

Total CPU time = 8.09 hours

Facit: there is always a way to improve on a previous estimation. We feel that also the 700 hours for the global model could be cut by the above procedure to a fraction of $1/3$ or better.

5. A proposal for a numerical solution of the free boundary value problem

5.1. Isoparametric elements.

Isoparametric elements are well known in finite element analysis. Confer e.g. Strang-Fix (1973), section 3.3. We outline the basic idea on hand of a two dimensional example. Consider the r, φ plane and consider an element partition as shown in figure 5.1(a). Denote by $S_{ij}(r, \varphi)$ the shape functions, where i refers to the node and j to a specific parameter at this node (Cf. section 3.2). We assume C^1 continuity of the $S_{ij}(r, \varphi)$.

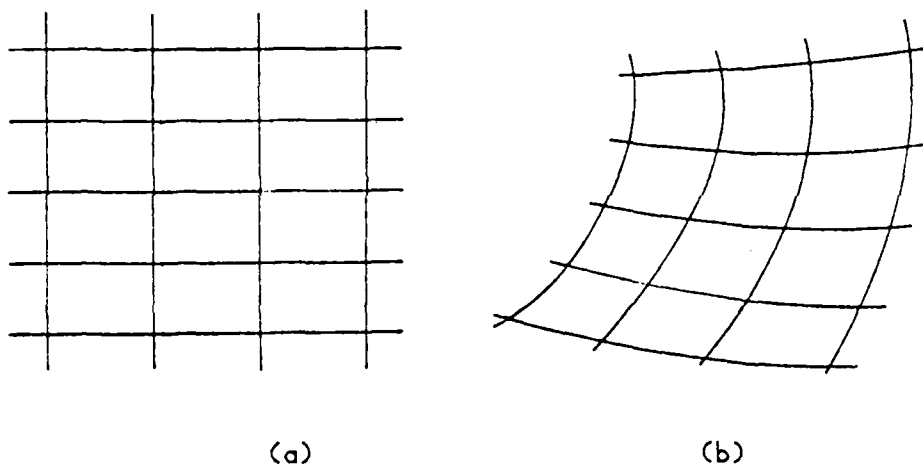


Figure 5.1

Mapping from straight to curved elements.

Consider now a mapping from the r, φ plane of figure 5.1(a) into the R, Φ plane of figure 5.1(b). The mapping functions are

$$\begin{aligned} R &= R(r, \varphi) \\ \Phi &= \Phi(r, \varphi) \end{aligned} \tag{5.1}$$

The elements get distorted thereby. The distortions shall be locally smooth. The mapping functions shall be C^1 continuous. In fact, we represent them in terms of finitely many parameters by using exactly the shape functions $S_{ij}(r, \varphi)$:

$$\begin{aligned} R &= \sum_{ij} R_{ij} S_{ij}(r, \varphi) \\ \phi &= \sum_{ij} \phi_{ij} S_{ij}(r, \varphi) \end{aligned} \quad (5.2)$$

It is now the R, ϕ plane of figure 5.1(b) where our problem in terms of a differential equation of the field and in terms of observations on this field is posed. The r, φ plane of figure 5.1(a) serves only an auxiliary purpose during the calculations.

There is Laplace's equation for the field:

$$\Delta V(R, \phi) = \frac{\partial^2 V}{\partial R^2} + \frac{1}{R} \frac{\partial V}{\partial R} + \frac{1}{R^2} \frac{\partial^2 V}{\partial \phi^2} = 0 \quad (5.3)$$

And there are point measurements involving $V(R, \phi)$ and its derivatives.

In order to have something specific in mind, assume that the image of the unit circle $r = 1$ in the r, φ plane is the (finite element approximation of the) level surface of the reference potential in the R, ϕ plane. Assume that point gravity anomalies are prescribed at certain locations $k = 1, 2, \dots, K$ which are the images of $r = 1, \varphi = \varphi_k$. Let $\nu = \nu_k$ be the normal to the level surface in the R, ϕ plane. Then the following equations hold ($V \dots$ disturbing potential, $g \dots$ normal gravity):

$$\Delta g_k = - \frac{\partial V}{\partial \nu} + \frac{1}{r} \frac{\partial r}{\partial \nu} V \bigg|_{\substack{R = R(1, \varphi_k) \\ \phi = \phi(1, \varphi_k)}} \quad (5.4)$$

Our next problem is to represent the potential in the R, ϕ plane in terms of finitely many parameters. Here, a deliberate detour is taken. One starts representing the image of V in the r, φ plane in terms of the shape functions:

$$V(r, \varphi) = \sum_{ij} V_{ij} S_{ij}(r, \varphi) \quad (5.5)$$

The representation of $V(R, \phi)$ is then obtained by means of the inverse functions

$$\begin{aligned} r &= r(R, \phi) = R^{-1}(R, \phi) \\ \varphi &= \varphi(R, \phi) = \phi^{-1}(R, \phi) \end{aligned} \quad (5.6)$$

Fortunately enough, the inverse functions will never be used explicitly during the numerical calculations. We obtain

$$V(R, \phi) = \sum_{ij} V_{ij} S_{ij}(r(R, \phi), \varphi(R, \phi)) \quad (5.7)$$

Remark: The fact that the same shape functions S_{ij} are used in order to represent the transformation and the potential is responsible for the name "isoparametric elements".

We now outline the calculation of the field contribution to the normals. Inserting the above representation of $V(R, \phi)$ into the Laplacean (5.3), we obtain

$$\Delta V(R, \phi) = \sum_{ij} V_{ij} \left[\frac{\partial^2 S_{ij}}{\partial R^2} + \frac{1}{R} \frac{\partial S_{ij}}{\partial R} + \frac{1}{R^2} \frac{\partial^2 S_{ij}}{\partial \phi^2} \right] \quad (5.8)$$

The following manipulations serve to circumvent the explicit use of $r(R, \phi)$, $\varphi(R, \phi)$ during the subsequent calculations. Let us take the term

$$\frac{\partial^2 S_{ij}}{\partial R^2} \quad (5.9)$$

as an example. It holds that

$$\begin{aligned} \frac{\partial S_{ij}}{\partial R} &= \frac{\partial S_{ij}}{\partial r} \cdot \frac{\partial r}{\partial R} + \frac{\partial S_{ij}}{\partial \varphi} \cdot \frac{\partial \varphi}{\partial R} \\ \frac{\partial^2 S_{ij}}{\partial R^2} &= \frac{\partial^2 S_{ij}}{\partial r^2} \left(\frac{\partial r}{\partial R} \right)^2 + 2 \frac{\partial^2 S_{ij}}{\partial r \partial \varphi} \cdot \frac{\partial r}{\partial R} \cdot \frac{\partial \varphi}{\partial R} + \frac{\partial^2 S_{ij}}{\partial \varphi^2} \left(\frac{\partial \varphi}{\partial R} \right)^2 \quad (5.10) \\ &+ \frac{\partial S_{ij}}{\partial r} \cdot \frac{\partial^2 r}{\partial R^2} + \frac{\partial S_{ij}}{\partial \varphi} \cdot \frac{\partial^2 \varphi}{\partial R^2} \end{aligned}$$

The functions

$$\frac{\partial S_{ij}}{\partial r}, \frac{\partial S_{ij}}{\partial \varphi}, \frac{\partial^2 S_{ij}}{\partial r^2}, \dots \quad (5.11)$$

don't require any further treatment because the S_{ij} are simple expressions in r, φ , e.g. bicubic polynomials which are differentiated as shown in section 3.1. Of course, after differentiation the substitution $r = r(R, \phi), \varphi = \varphi(R, \phi)$ has to be imagined but this substitution will readily be undone during a subsequent change of variables in the integrals. Only the derivatives of r, φ with respect to R cause trouble. One forms the Jacobian matrix of partials

$$\begin{bmatrix} \frac{\partial R}{\partial r} & \frac{\partial R}{\partial \varphi} \\ \frac{\partial \phi}{\partial r} & \frac{\partial \phi}{\partial \varphi} \end{bmatrix} = J(r, \varphi) \quad (5.12)$$

One inverts this matrix:

$$\begin{bmatrix} \frac{\partial r}{\partial R} & \frac{\partial r}{\partial \phi} \\ \frac{\partial \varphi}{\partial R} & \frac{\partial \varphi}{\partial \phi} \end{bmatrix} = J^{-1}(r, \varphi) \quad (5.13)$$

obtaining in this way the first order partials

$$\frac{\partial r}{\partial R}, \frac{\partial \varphi}{\partial R}, \dots \quad (5.14)$$

whereby the substitution $r = r(R, \phi)$, $\varphi = \varphi(R, \phi)$ must be imagined, but is not done explicitly. Denote temporarily

$$\begin{aligned} \frac{\partial r}{\partial R} &= e(r, \varphi) & \frac{\partial r}{\partial \phi} &= f(r, \varphi) \\ \frac{\partial \varphi}{\partial R} &= g(r, \varphi) & \frac{\partial \varphi}{\partial \phi} &= h(r, \varphi) \end{aligned} \quad (5.15)$$

Then

$$\begin{aligned} \frac{\partial^2 r}{\partial R^2} &= \frac{\partial e(r, \varphi)}{\partial r} \cdot \frac{\partial r}{\partial R} + \frac{\partial e(r, \varphi)}{\partial \varphi} \cdot \frac{\partial \varphi}{\partial R} = \\ &= \frac{\partial e(r, \varphi)}{\partial r} \cdot e(r, \varphi) + \frac{\partial e(r, \varphi)}{\partial \varphi} g(r, \varphi) \end{aligned} \quad (5.16)$$

similarly:

$$\frac{\partial^2 r}{\partial \phi^2} = \frac{\partial f(r, \varphi)}{\partial r} \cdot f(r, \varphi) + \frac{\partial f(r, \varphi)}{\partial \varphi} h(r, \varphi) \quad (5.17)$$

Hence the second derivatives needed in

$$\frac{\partial^2 S_{ij}}{\partial R^2} \quad (5.18)$$

are calculated. In an analogous fashion all quantities needed in evaluating $\Delta V(R, \phi)$ can be obtained. One arrives at the following general expression

$$\begin{aligned} \Delta V(R, \phi) &= \sum_{ij} V_{ij} \Delta S_{ij}(r(R, \phi), \varphi(R, \phi)) = \\ &= \sum_{ij} V_{ij} Q_{ij}(r(R, \phi), \varphi(R, \phi)) \end{aligned} \quad (5.19)$$

where the substitution $r = r(R, \phi)$, $\varphi = \varphi(R, \phi)$ is not actually carried out.

We proceed to the integration over the curved element in the R, ϕ plane:

$$\begin{aligned} E &= \frac{1}{2} \int_{\Omega} (\Delta V(R, \phi))^2 R dR d\phi \\ &= \frac{1}{2} \int_{\Omega} \left[\sum_{ij} V_{ij} Q_{ij}(r(R, \phi), \varphi(R, \phi)) \right]^2 R dR d\phi \end{aligned} \quad (5.20)$$

Taking the variation with respect to the V_{ij} , one gets (as in section 3.2.2.) the contribution from this quad to the normals as

$$\sum_{ij} g_{ij; i'j'} V_{i'j'} = 0 \quad (5.21)$$

with

$$g_{ij; i'j'} = \int_{\Omega} Q_{ij}(r(R, \phi), \varphi(R, \phi)) \cdot Q_{i'j'}(r(R, \phi), \varphi(R, \phi)) R dR d\phi \quad (5.22)$$

Now comes the announced transformation of parameters of the r, φ plane. The domain is mapped onto the uncurved rectangle in the r, φ plane. $\text{Det}[J(r, \varphi)]$ is the Jacobian determinant of this transformation.

$$g_{ij}; i, j' = \int_{\Omega} Q_{ij}(r, \varphi) Q_{ij'}(r, \varphi) R(r, \varphi) \cdot \text{Det}[J(r, \varphi)] dr d\varphi \quad (5.23)$$

Remark: Although the functions $Q_{ij}(r, \varphi)$, $R(r, \varphi)$ readily split into sums of products of terms $F(r)G(\varphi)$, the Jacobian $\text{Det}[J(r, \varphi)]$ does not. Hence the integration procedure is not easily carried out. Numerical integration may be necessary.

Let us also elaborate on the contribution of the measurements, in our special case, the gravity anomalies. First we need a representation of the vector ν , i.e. the normal to the level surface

$$\begin{aligned} R &= R(1, \varphi) \\ \phi &= \phi(1, \varphi) \end{aligned} \quad (5.24)$$

We may view this as a parameter representation of the surface. In the orthonormal system

$$e_R = \frac{\partial}{\partial R}, \quad e_\phi = \frac{1}{R} \cdot \frac{\partial}{\partial \phi} \quad (5.25)$$

the vector tangent to the level surface (curve) has the representation

$$\begin{bmatrix} \frac{\partial R}{\partial \varphi} \\ R \frac{\partial \phi}{\partial \varphi} \end{bmatrix} \quad (5.26)$$

Hence the orthogonal vector has the representation

$$\begin{bmatrix} R \frac{\partial \phi}{\partial \psi} \\ -\frac{\partial R}{\partial \psi} \end{bmatrix} \quad (5.27)$$

Making this a unit vector

$$\frac{1}{\sqrt{\left(\frac{\partial R}{\partial \psi}\right)^2 + \left(R \frac{\partial \phi}{\partial \psi}\right)^2}} \begin{bmatrix} R \frac{\partial \phi}{\partial \psi} \\ -\frac{\partial R}{\partial \psi} \end{bmatrix} \quad (5.28)$$

Hence

$$\frac{\partial V}{\partial \psi} = \frac{1}{\sqrt{\left(\frac{\partial R}{\partial \psi}\right)^2 + \left(R \frac{\partial \phi}{\partial \psi}\right)^2}} \left[\frac{\partial V}{\partial R} \cdot R \frac{\partial \phi}{\partial \psi} - \frac{\partial V}{\partial \phi} \cdot \frac{\partial R}{\partial \psi} \right] \quad (5.29)$$

Again

$$\begin{aligned} \frac{\partial V}{\partial R} &= \frac{\partial V}{\partial r} \cdot \frac{\partial r}{\partial R} + \frac{\partial V}{\partial \psi} \cdot \frac{\partial \psi}{\partial R} \\ \frac{\partial V}{\partial \phi} &= \frac{\partial V}{\partial r} \cdot \frac{\partial r}{\partial \phi} + \frac{\partial V}{\partial \psi} \cdot \frac{\partial \psi}{\partial \phi} \end{aligned} \quad (5.30)$$

where the formulas (5.13) are employed once more.

Discussion: Remember that we are presently dealing with a completely known isoparametric transformation. The parameters R_{ij} , ϕ_{ij} are prespecified. What is the benefit from such a transformation and what is the price to be paid?

An isoparametric transformation opens the way to exotically shaped reference surfaces. The resulting normal equations are still as sparse as those for the spherical (or ellipsoidal) reference surface. Hence the solution of the normals require no additional effort. A price has to be paid when the normals are formed. The integration procedure is more costly with respect to programming effort and computation time. On the other hand, the solution of the normals is asymptotically more time consuming than the formation. If N is the No. of elements at ground level, then the formation of normals is an effort of $c_f N$ where on the solution is an effort $c_s N^{\frac{1}{2}}$. Here c_f and c_s are (nearly) constants. Recall, by the way, that a solution by means of surface layer elements or collocation requires cN^3 ! It is seen that isoparametric elements affect the constant factor c_f but not c_s .

5.2. Approaching the free boundary value problem of physical geodesy

Isoparametric elements as outlined in the previous section open a door to approach the fundamental problem of physical geodesy in a more direct way than this has been done thus far. Our subsequent presentation will be expository. Detailed formulas will be given elsewhere.

The fundamental problem of physical geodesy is the simultaneous determination of the earth's figure and potential from measurements of the 3-dimensional gravity vector at the earth's surface. The historical approach to this problem seemingly has been done under the motto: "First linearize everything in sight and then think about formulating a meaningful problem". As already proposed in Meissl (1971) we prefer to formulate a problem and then to linearize it. However, we assume that the unknown surface of the earth is smooth. This means that the terrain has been smoothed and that the gravity measurements have been corrected accordingly.

Assume that a sphere of radius r_0 situated in r, φ, λ space is the pre-image of the earth's surface under a certain unknown transformation

$$\begin{aligned} R &= R(r, \varphi, \lambda) \\ \Phi &= \Phi(r, \varphi, \lambda) \\ \Lambda &= \Lambda(r, \varphi, \lambda) \end{aligned} \quad (5.31)$$

Assume that the measurements g_1, g_2, \dots of the 3 dimensional gravity vector are taken at locations identified by their coordinates $r_k = r_0, \varphi_k, \lambda_k$ in r, φ, λ space. The vectors g_k themselves are considered to be represented in a rectangular equatorial system. Assume further, that the earth's rotation is known so that the rotational part of the potential can be taken into account by a known reference potential U , while the unknown disturbing potential V can be assumed harmonic.

The transformation from r, φ, λ space to R, Φ, Λ space is now set up according to equations (5.2) in the previous section as

$$\begin{aligned} R &= \sum_{ij} R_{ij} S_{ij}(r, \varphi, \lambda) \\ \Phi &= \sum_{ij} \Phi_{ij} S_{ij}(r, \varphi, \lambda) \\ \Lambda &= \sum_{ij} \Lambda_{ij} S_{ij}(r, \varphi, \lambda) \end{aligned} \quad (5.32)$$

The $R_{ij}, \Phi_{ij}, \Lambda_{ij}$ are assumed known for all nodes i except those situated at the earth's surface. We define the index set L in a way that $i \in L$ identifies precisely these nodes. Thus $R_{ij}, \Phi_{ij}, \Lambda_{ij}$ are assumed unknown for $i \in L$. The remaining coefficients $R_{ij}, \Phi_{ij}, \Lambda_{ij}, i \notin L$ are known and can be chosen according to computational convenience. For example they can establish a transition from spherical surfaces $r = \text{const}$ to ellipsoidal surfaces in the R, Φ, Λ space.

Next the contribution of the field to the normal equations is considered. One forms the energy integral

$$E = \frac{1}{2} \int_{\Omega} [\Delta V(R, \Phi, \Lambda)]^2 R \cos \Phi \, dR \, d\Phi \, d\Lambda \quad (5.33)$$

It is now extended over the entire exterior space in R, Φ, Λ space. Unknown quantities in this integral are not only the nodal values V_{ij} but also $R_{ij}, \Phi_{ij}, \Lambda_{ij}, i \in L$. Hence variation of E must be performed with respect to all these quantities. Thereby the usual decomposition of the domain into the individual elements is employed, and a transformation of the element integrals back to r, φ, λ space can be done just as outlined in the previous section. The resulting normal equations are linear with respect to V_{ij} but nonlinear with respect to $R_{ij}, \Phi_{ij}, \Lambda_{ij}, i \in L$. They are of the following form

$$\sum_{ij} g_{ij; i'j'} (R_{ke}, \Phi_{ke}, \Lambda_{ke}) V_{i'j'} = 0 \quad (5.34)$$

The equations are still sparse if considered as linear equations in V_{ij} . However sparseness also extends to the nonlinear equations in the following sense. Any nonzero coefficient $g_{ij; i'j'}$ is related to one or more elements coupling the nodes i and i' . Hence $g_{ij; i'j'}$ will only involve such transformation parameters $R_{ke}, \Phi_{ke}, \Lambda_{ke}, k \in L$ which belong to these common elements. In other words, if the nonlinear normal equations are linearized, then the resulting equations will have the same sparsity pattern as the equations of the earlier chapters. The only difference is that all nodes $i \in L$ will have additional parameters $R_{ij}, \Phi_{ij}, \Lambda_{ij}$.

The contribution to the normals from the assumed gravity measurements does not pose much difficulty. The observation equations

are of the following type

$$\begin{aligned}\frac{\partial}{\partial x} \left\{ U(R, \phi, \lambda) + V(R, \phi, \lambda) \right\}_k &= g_k^{(x)} + r_k^{(x)} \\ \frac{\partial}{\partial y} \left\{ U(R, \phi, \lambda) + V(R, \phi, \lambda) \right\}_k &= g_k^{(y)} + r_k^{(y)} \\ \frac{\partial}{\partial z} \left\{ U(R, \phi, \lambda) + V(R, \phi, \lambda) \right\}_k &= g_k^{(z)} + r_k^{(z)}\end{aligned}\quad (5.35)$$

U is the reference potential, V the disturbing potential. The equations are again linear in V_{ij} , but nonlinear in $R_{ij}, \phi_{ij}, \lambda_{ij}, i \in L$. The reference potential U also gives a contribution due to the unknown location of the measurements, i.e. due to the unknowns $R_{ij}, \phi_{ij}, \lambda_{ij}, i \in L$. The residuals $r_k^{(x)}, r_k^{(y)}, r_k^{(z)}$ are weighted and nonlinear normals are formed. They will show an analogous sparsity pattern after linearization.

The normals of field and measurements are added and solved by Newton's method. This method requires precisely that linearization which we were talking about above.

We shall conclude this outline of an intended research project with the following

Remark: It is not considered that the use of three-dimensional gravity measurements is very meaningful in the geodetic boundary value problem unless additional information on horizontal position is available. Such information comes either from ground control networks or from space observations. Hence one should attempt to either incorporate this information (without destroying the sparsity pattern) or to remove the horizontal degrees of freedom in the unknown mapping from r, φ, λ space to R, ϕ, λ space.

6. Computer experiments for 2-dimensional problems

6.1. Purpose and scope

The computer experiments to be described in this chapter were designed to find an answer to the following specific questions.

- 1) Should the finite element approach be based on the Ritz-, or the Trefftz-, or on the least squares principle. As we have pointed out in section 3.3.1. the least squares principle was found to perform best in the presence of noisy and redundant data.
- 2) Are cubic polynomials sufficiently accurate, or should quintic polynomials be used? The question is perhaps posed in an overly simplified way. One can always account for the lower degree of the cubics by choosing a smaller element partition. Hence the question should be asked as follows. Does it pay off to replace the choice of cubics and an adequate element partition by the use of quintics and an element partition having appropriately larger elements? Quintics look attractive because they are C^2 continuous. One can even enforce the Laplacean to vanish at the nodes. Nevertheless it was found that the use of cubics is preferable.
- 3) Can the attenuation-with-altitude-effect of the potential be exploited in a way that the size of the elements increases with altitude fast enough to ensure that the total number of nodes is bounded by a constant times the number of nodes at the surface. The answer is affirmative.
- 4) What is the best way to represent the field in the remote outer space of the earth? It is believed that specially designed elements of infinite size together with appropriately chosen local shape functions

(different from cubic polynomials) are the best choice. Confer section 3.6.

5) What is the best ratio between weights applied to the field contribution to the normals and those applied to the contribution from the geodetic observations?

6) Is there a way to combine in one calculation surface data with satellite derived spherical harmonics?

Due to the large CPU-times predicted for 3-dimensional calculations it was decided to conduct the experiments in 2 dimensions. Stoke's problem for the unit circle was solved by means of finite elements, and for a set of artificially generated data. Our version of Stoke's problem is formulated as follows. Find a potential $V(r, \varphi)$ in the outer space of the unit circle such that

$$(1) \quad V(r, \infty) = O\left(\frac{1}{r^2}\right) ; \quad r \rightarrow \infty \quad (6.1)$$

$$(2) \quad \Delta V = 0 ; \quad r > 1 \quad (6.2)$$

$$(3) \quad V(r, \varphi) + \frac{\partial V}{\partial r} = f(\varphi) ; \quad r = 1 \quad (6.3)$$

It is seen that we have eliminated the logarithmic part of the potential. V may be viewed as a disturbing potential. The reference potential U may absorb the logarithmic part. The function $f(\varphi)$ must be free of "circular harmonics" of the zero-th and first degree. (I.e. its Fourier series must start with terms in $\cos 2\varphi$, $\sin 2\varphi$).

Remark: Obviously the stated problem is most easily solved by means of circular harmonics, i.e. by Fourier-analysis. Representing $f(\varphi)$ as

$$f(\varphi) = \sum_{n=2}^{\infty} (c_n \cos n\varphi + d_n \sin n\varphi) \quad (6.3a)$$

the solution is obtained as

$$V(r, \varphi) = \sum_{n=2}^{\infty} \frac{1}{r^n} (a_n \cos n\varphi + b_n \sin n\varphi) \quad (6.3b)$$

with

$$a_n = -\frac{c_n}{n-1} \quad b_n = -\frac{d_n}{n-1} \quad (6.3c)$$

The above continuous version (6.1-3) of Stoke's problem is replaced by a discrete one. First, the continuous data $f(\varphi)$ are replaced by $f(\varphi_k)$ for discrete arguments φ_k . Secondly, the potential $V(r, \varphi)$ is replaced by a finite element representation

$$\sum_{ij} V_{ij} S_{ij}(r, \varphi) \quad (6.4)$$

Confer sections 3.2. and 3.3.1, in particular equation (3.19). The element partition is shown in principle in figures 3.6(a)-(b) of section 3.3.2. In all experiments conducted thus far, the arguments φ_k were assumed equally spaced.

The outer zone was represented in two ways, namely by a "circular harmonics" representation of comparatively low degree (cf. section 3.7, method labeled (4)), and alternatively by elements of infinite size as outlined in section 3.6.

The data $f(\varphi_k)$ were generated from an "assumed potential" of the form

$$V_{\text{ASS}}(r, \varphi) = \sum_{n=2}^{N_{\text{MAX}}} \frac{1}{r^n} \{ a_n \cos n\varphi + b_n \sin n\varphi \} \quad (6.5)$$

with $a_n = \frac{\alpha_n}{n^d} \quad , \quad b_n = \frac{\beta_n}{n^d}$

The coefficients α_n, β_n were randomly generated in the interval $[-0.5, +0.5]$. The maximal degree N_{\max} , and the damping exponent d were varied. The choice $d = 1$ corresponds to Kaula's rule of the thumb. (Equal degree variances of the radial derivative for a wide range $2 \leq n \leq N_{\max}$).

Remark: Note that no measurement noise was assumed to be superimposed upon the data calculated from the assumed potential.

The procedure described in chapter 3 and section 4.1 yielded the potential at the nodes of the chosen element partition together with the nodal derivatives. This "calculated potential" was compared with the assumed potential (6.5). Statistics of the deviations were calculated and tabulated. The calculations were carried out on the OSU-Computer Amdahl 470 V/6-II. In a few cases, a post-analysis of the OSU results was done on a desk-top computer WANG 2200 VP. The potential was then interpolated into the interior of the finite elements in order to verify that the approximation was also good there. Also a few orbits of passive masspoints were numerically integrated for the assumed potential and, alternatively, for the calculated potential. In both cases a circular symmetric reference potential of the $\log r$ type was superimposed.

6.2. Parameters distinguishing the experiments.

In this section we give a detailed description of the input parameters characterizing one particular experiment.

IDEGR	degree of polynomials in elements of finite size:
3	... bicubic polynomials
5	... biquintic polynomials

ISWINF a switch distinguishing the representation of the field in
the remote outer zone $r > r_{out}$
1 ... circular harmonics
2 ... elements of infinite size

NLAYER number of successive layers of elements in the circular
ring $1 \leq r \leq r_{out}$

ITYPE(I), $I = 1, \dots, NLAYER$ type of elements in layer number I
1 ... simple quads
2 ... compound quads

NELEM(1) number of elements in layer number 1 (the lowermost)
layer

HFCT a factor governing the thickness of the various layers
and also responsible for the size of r_{out} . Cf. the following
remark.

Remark: From the input parameters NLAYER, ITYPE(I), $I = 1, \dots, NLAYER$,
NELEM(1), HFCT the element partition was calculated by the following set
of formulas:

$$NELEM(I) = NELEM(I-1)/ITYPE(I-1); \quad I = 2, \dots, NLAYER \quad (6.6)$$

$$DELPHI(I) = 2 * \pi / NELEM(I); \quad I = 1, \dots, NLAYER \quad (6.7)$$

$$RADIUS(1) = 1$$

$$RADIUS(I) = RADIUS(I-1) * (1 + DELPHI(I) * HFCT) \\ I = 2, \dots, NLAYER+1 \quad (6.8)$$

Thereby we have denoted

NELEM(I) number of elements in layer I

DELPHI(I) angular width of elements in layer I

RADIUS(I) inner radius of layer I
 outer radius of layer I-1
 I = 1, ..., NLAYER

RADIUS(NLAYER+1) = r_{out} outer radius of last layer = outer radius
 of the circular ring partitioned into elements of finite
 size.

We now continue to describe the input parameters.

NPOTC largest degree of circular harmonics in assumed potential
 (denoted N_{max} in equation (6.5))

DAMP damping factor in the assumed potential (denoted d in
 equation (6.5))

NPINST this parameter governs the number of locations at which
 data (gravity anomalies) were calculated from the
 assumed potential. Remember that the locations are equally
 spaced. NPINST is the number of locations in an interval
 of size DELPHI(1), i.e. in a boundary segment of a
 lowermost finite element at $r = 1$. The data are
 arranged there as shown in figure 6.1. It is seen that
 the interval ends corresponding to element boundaries
 are halfway situated between two measurements locations.

The total number of fictitious measurements is thus obtained as $NELEM(1)*NPINST$. Recall that no simulated measurement noise was superimposed upon the data.

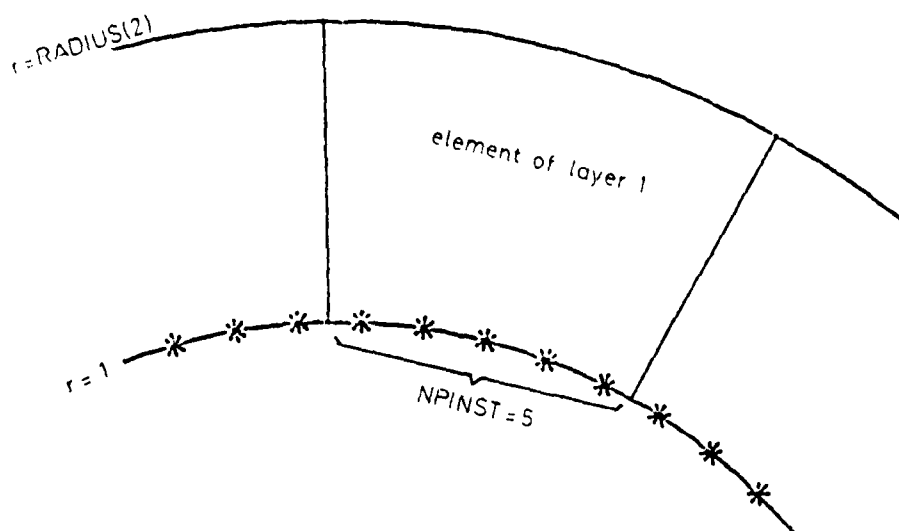


Figure 6.1. Arrangement of fictitious measurements in a boundary segment of the unit sphere

STKWGT weight applied to the contribution of the measurements to the normals. (The weight for the field contribution was assumed with a value equal to 1).

The following 4 parameters apply only to the case $ISWINF = 1$, i.e. to the representation of the field in $r > r_{out}$ by circular harmonics.

NHARM highest degree of circular harmonics in the representation of the field in the outer zone. This representation is thus given by

$$V(r, \varphi) = \sum_{n=2}^{NHARM} \frac{1}{r^n} \{ A_n \cos n\varphi + B_n \sin n\varphi \} \quad (6.9)$$

The calculated values for the unknowns A_n , B_n should approximate those of a_n , b_n in the representation (6.5) of the assumed potential. The deviations are due to

- (1) discretization of data
- (2) discretization of the field by the finite element representation
- (3) $NHARM \leq NPOTC$
- (4) roundoff error

NPINHA

this parameter governs the number of locations at which the finite element representation of the field in the ring was collocated with the circular harmonics representation of equation (6.9) in the outer zone. The locations are all found at $r = r_{out}$. They are equally spaced. There are $NELEM(NLAYER)$ intervals at $r = r_{out}$. NPINHA gives the number of locations in one of these intervals. Their distribution is similar as that shown in figure 6.1, i.e. the locations are equally spaced and the interval boundaries were assumed halfway between the two adjacent locations. The total number of points of collocation is thus $NELEM(NLAYER)*NPINHA$. This number was frequently assumed appreciably larger than $2*NHARM - 2$, the number of coefficients in (6.9). This means that we are working with a redundant set of locations at which consistency of the two potentials was enforced. Hence remark 3 given in section 3.7 applies mutatis mutandis (i.e. translated to the method labeled (4) in section 3.7). Consistency can only be required in a least squares sense. Weight assumptions were necessary. They are specified by the subsequent 2 weight parameters.

HARWG0 weight applied to the values of the potential during collocation at the locations on the circle $r = r_{out}$

HARWG1 weight applied to the radial derivative of the potential during collocation.

Remark: Recall that the weight applied to the field contribution was assumed equal to 1.

The following 3 parameters apply only to the case of IDEGR = 5, i.e. to the case of biquintics.

N89 This parameter allows to choose between 8 parameters per node and 9 parameters per node. In case of 8 parameters per node, the Laplacean is fixed to zero at any node. This gives a linear relation between the nodal parameters V_{rr} , V_r , $V_{\theta\theta}$ which was used in order to eliminate V_{rr} .

FLPWGL Because it was observed that the nodal parameters involving second radial derivatives were rather poorly determined at $r = 1$, additional fictitious observations of the Laplacean were assumed at all measurement locations. The appropriate weight was FLPWGL.

FLPWGU This input parameter is analogous to the previous one, however it applies to $r = r_{out}$.

6.3. Detailed results for two experiments.

6.3.1. An experiment using comparatively large elements.

The input parameters were specified as follows

IDEGR	=	3	[cubics]
ISWINF	=	2	[circular harmonics in outer zone]
NLAYER	=	4	[number of layers]
ITYPE(1)	=	1	[simple quads at first layer]
ITYPE(2)	=	1	[simple quads at second layer]
ITYPE(3)	=	2	[compound quads at third layer]
ITYPE(4)	=	2	[compound quads at fourth layer]
NELEM(1)	=	32	[number of elements in first layer]
HFCT	=	1	[approximately square shaped elements]
NPOTC	=	128	[highest degree in assumed potential]
DAMP	=	2	[damping factor in assumed potential]
NPINST	=	8	[8 data points per interval at $r = 1$]
STKWGT	=	5	[weight for data]
NHARM	=	8	[highest degree of circular harmonics in the representation of the field in the outer zone]
NPINHA	=	8	[8 collocation points per interval at $r = r_{out}$]
HARWG0	=	50	[weight in collocating V at $r = r_{out}$]
HARWG1	=	5	[weight in collocating V_r at $r = r_{out}$]

Table 6.1 illustrates the geometry of the element partition. A pictorial representation is given by our earlier figures 3.6(a), (b) in section 3.3.2.

LAYER	ITYPE	NELEM	RADIUS	DELPHI
1	1	32	1.00000	0.19635
2	1	32	1.19635	0.19635
3	2	16	1.43125	0.39270
4	2	8	1.99330	0.78540

$$r_{out} = 3.55884$$

Table 6.1. Element partition in experiment
with comparatively large elements

The assumed potential (equation (6.5)) is depicted at ground level ($r = 1$) in figure 6.2(a). It is seen that the potential does not exceed the value 0.14. Recall that the assumed potential was used to generate the discrete set of gravity anomalies. From these data the potential was calculated backwards by the finite element method. The calculated potential at $r = 1$ is shown in figure 6.2(b). It is seen that the irregularities are somewhat smoothed out. Figure 6.2(c) shows the difference. Note that the ordinates are now scaled differently. We see that the relative error is about 2.5%.

The figure 6.3(a)-(c) describe in a similar way the behaviour of the radial derivative V_r at $r = 1$. Figures 6.4(a)-(c) are devoted to the horizontal derivative V_y at $r = 1$.

It is obvious that the accuracy in the presently described experiment is insufficient. In order to obtain an approximation of the geoid at the cm level, the finite element calculation should reproduce about 4 correct digits of the disturbing potential V . Presently only 2 digits are correct. We have nevertheless exhibited the results in some detail because they tell us very instructively what we can expect qualitatively from a finite element solution. If the elements are chosen too large, the details of the field can not be properly represented. One should nevertheless expect that the approximation is good in the low

frequencies. This is apparently the case and is further illustrated by figures 6.2(d-e), 6.3(d-e), 6.5 and 6.6.

Figure 6.2(d) shows the assumed field V at $r = 1$, truncated to circular harmonics of degree $n \leq 32$. This low-frequency part of the field has about the same number of parameters as the trace of the finite element representation of V at $r = 1$. The two graphs 6.2(b) and 6.2(d) are visually nearly indistinguishable. Figure 6.2(e) shows the high frequency part of the assumed field, composed of circular harmonics of degree $32 < n \leq 128$. The two graphs of figures 6.2(c) and 6.2(e) are different, but the magnitude is the same. Figures 6.3(d-e), which should be compared to figures 6.3(b-c) suggest that the same conclusion holds for the radial derivative: the finite element solution is about as good as the assumed field truncated to low-degree harmonics of degree $n \leq 32$.

Figure 6.5 shows the superposition of figures 6.3(a) and 6.3(b). Figure 6.6 shows the calculated Laplacean ΔV at $r = 1$. (The assumed Laplacean is zero, of course). The calculated Laplacean is discontinuous at element boundaries. Hermite bicubics are only C^1 continuous. The graph of ΔV suggests that the finite element solution achieves a smoothing by shifting - in a balanced way - positive and negative masses outward of the earth's body. The method automatically regularizes the field in this way. The shifting of masses should not be noticeable at the low frequencies.

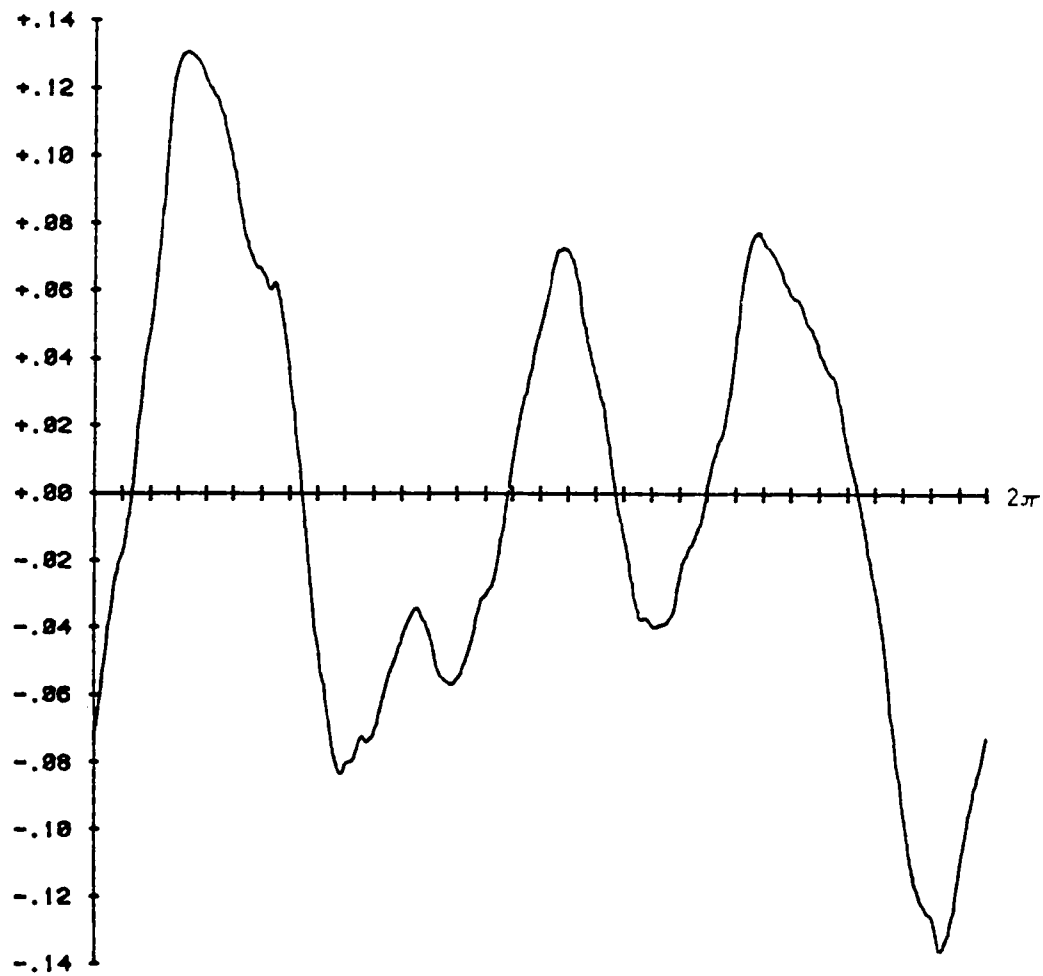


Figure 6.2(a). Assumed potential V at $r = 1$ in experiment using comparatively large elements.

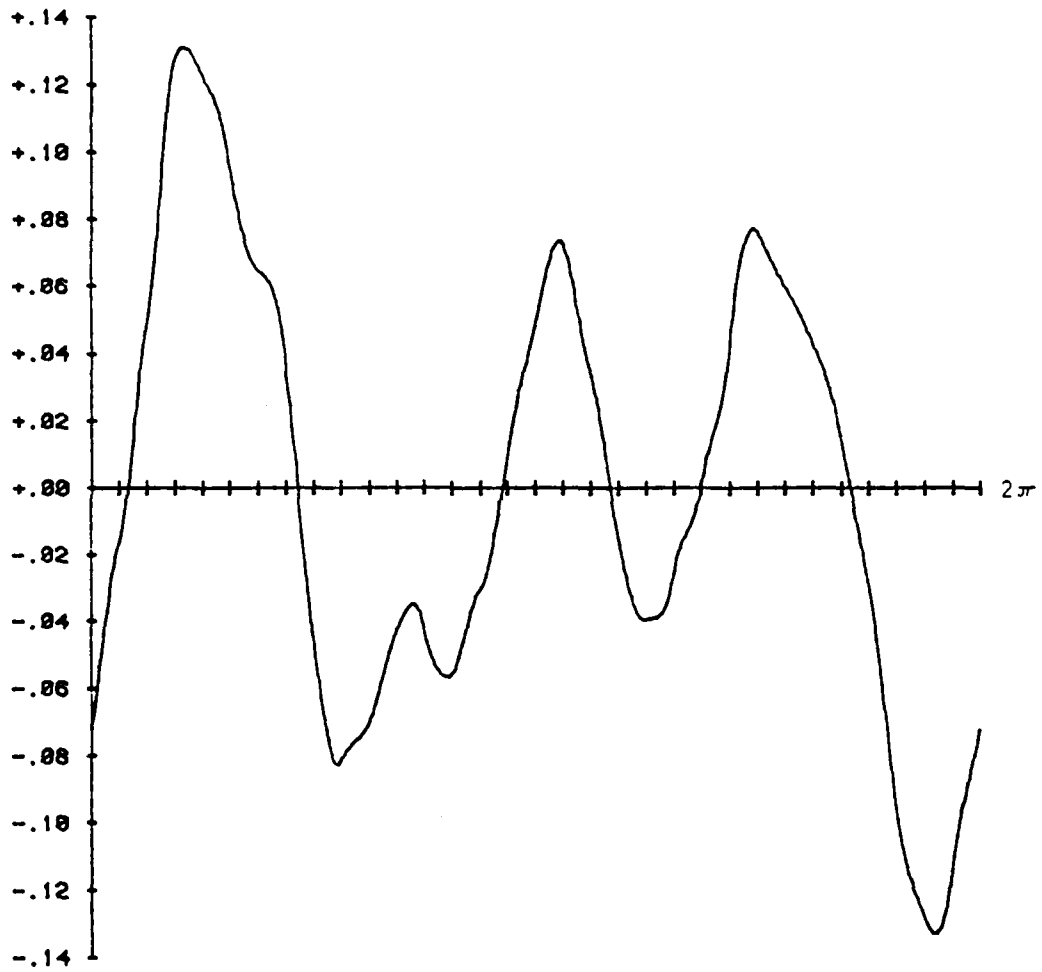


Figure 6.2(b). Calculated potential V at $r = 1$ in experiment using comparatively large elements.

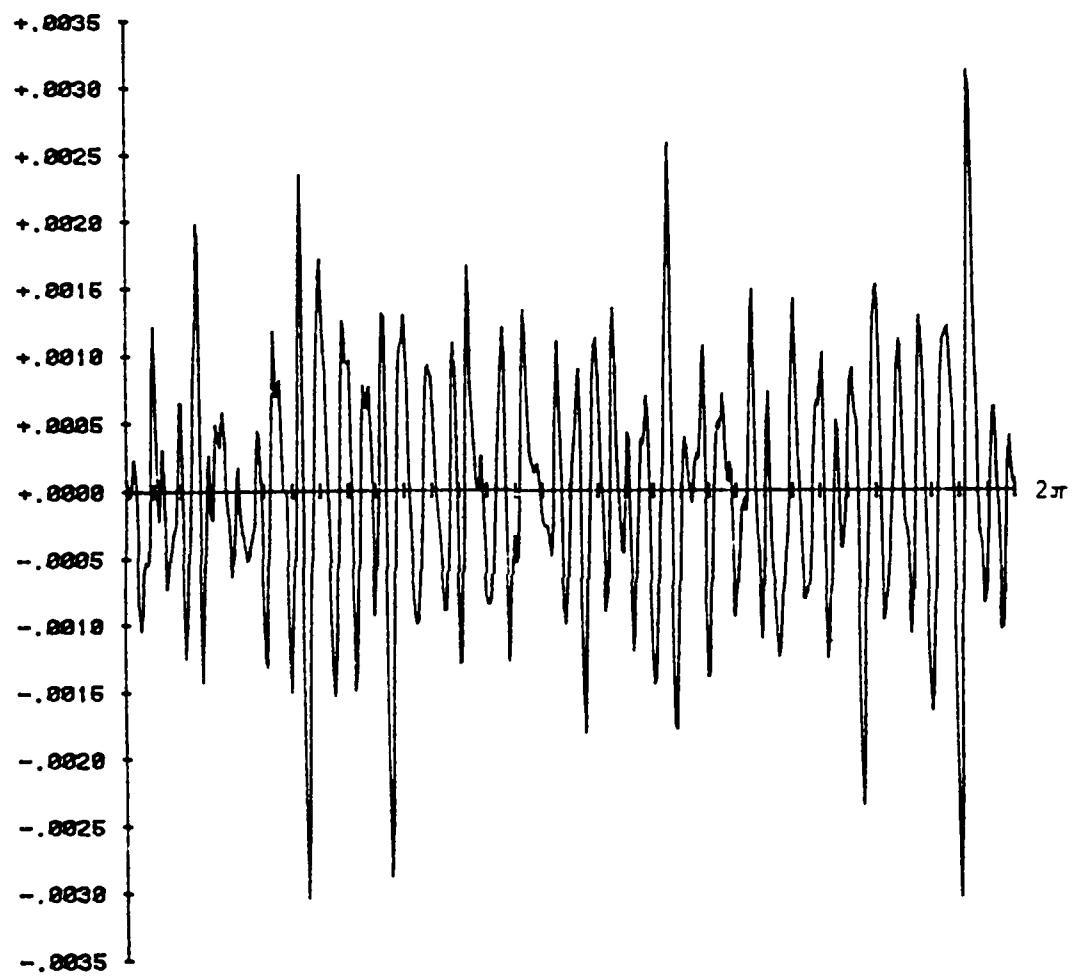


Figure 6.2(c). Difference calculated minus assumed potential V at $r = 1$ in experiment using comparatively large elements.

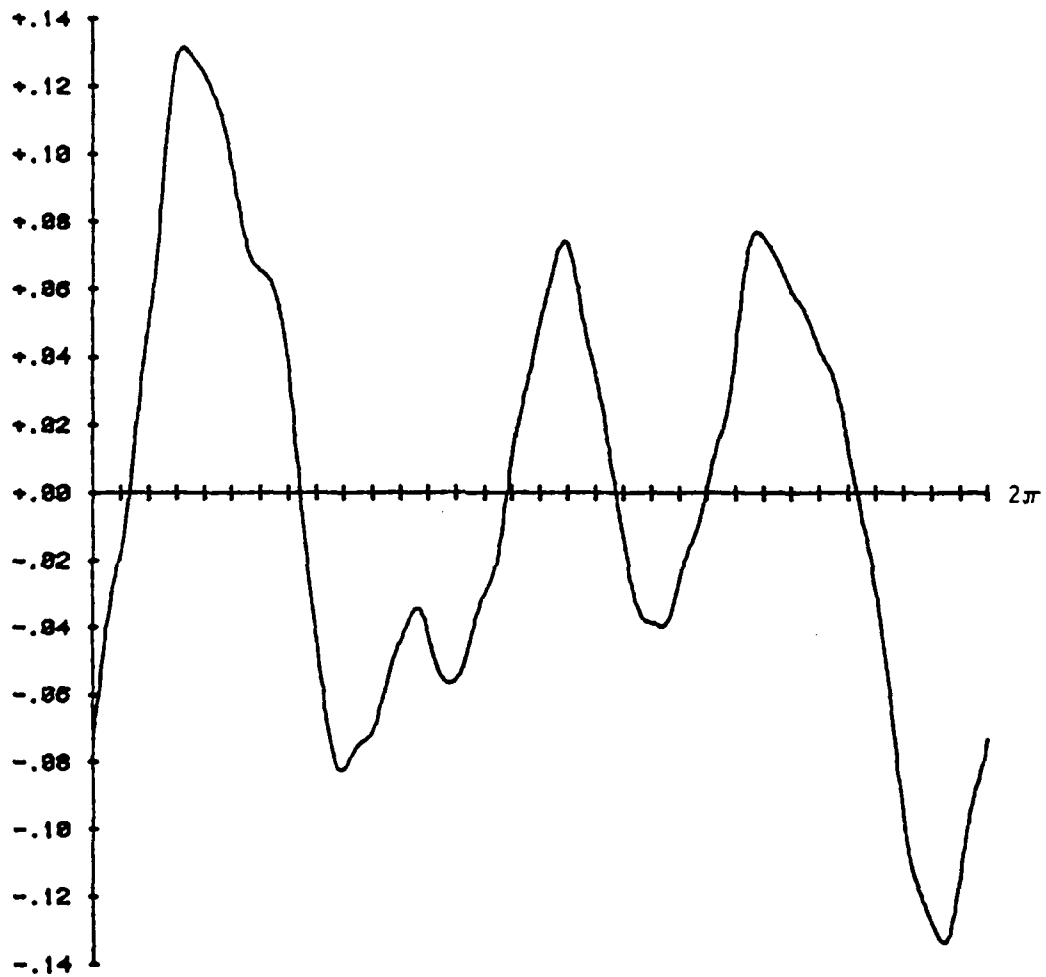


Figure 6.2(d). The assumed field V at $r = 1$ truncated to circular harmonics of degree $n < 32$. Experiment using comparatively large elements.

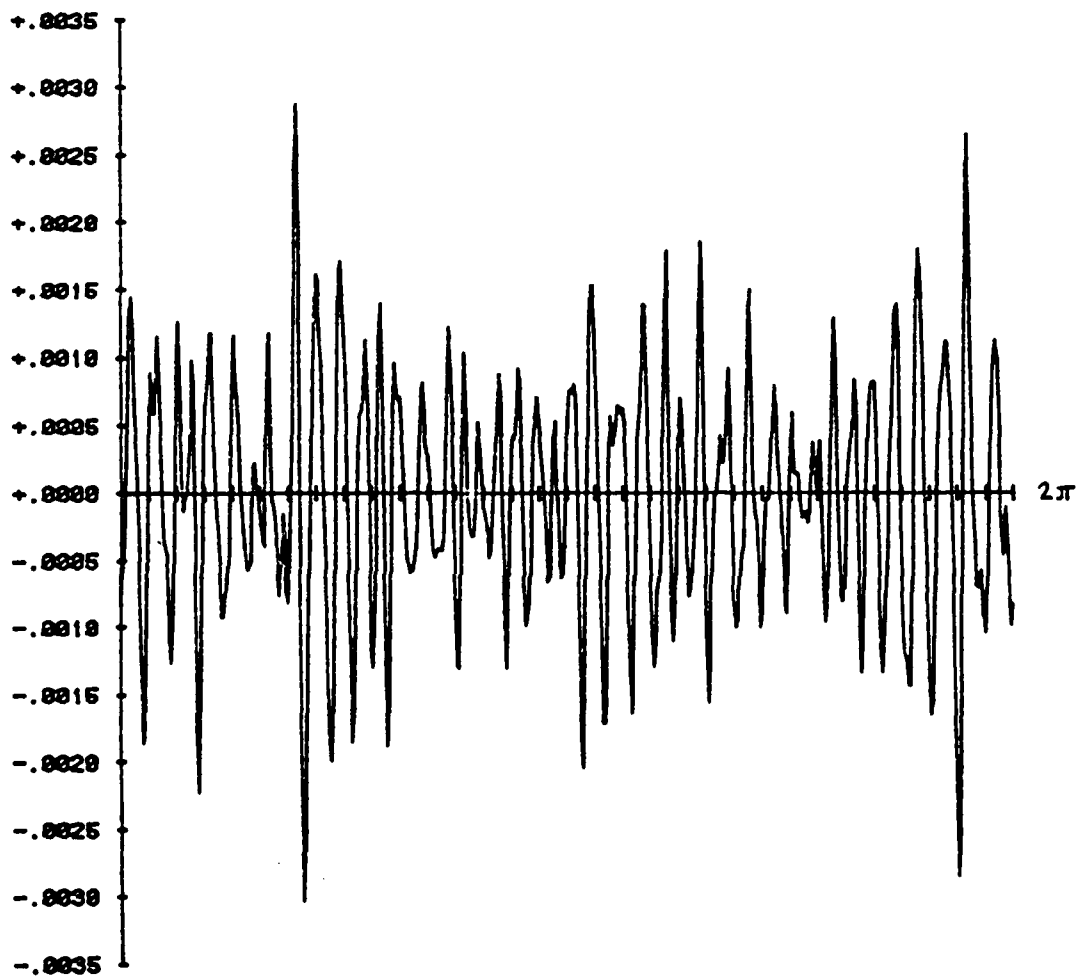


Figure 6.2(e). Contribution of circular harmonics of degree $n > 32$ to the assumed field V at $r = 1$. Experiment using comparatively large elements.

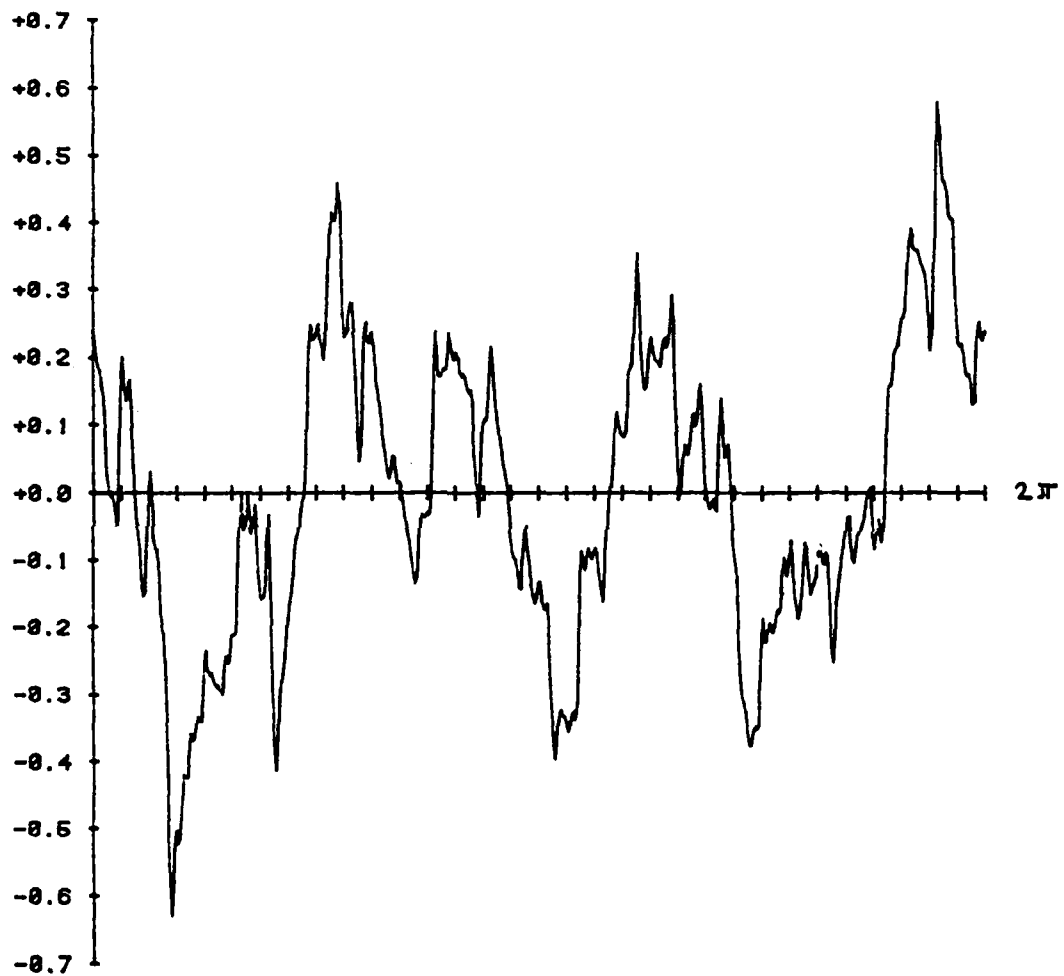


Figure 6.3(a). Assumed radial derivative V_r at $r = 1$ in experiment using comparatively large elements.

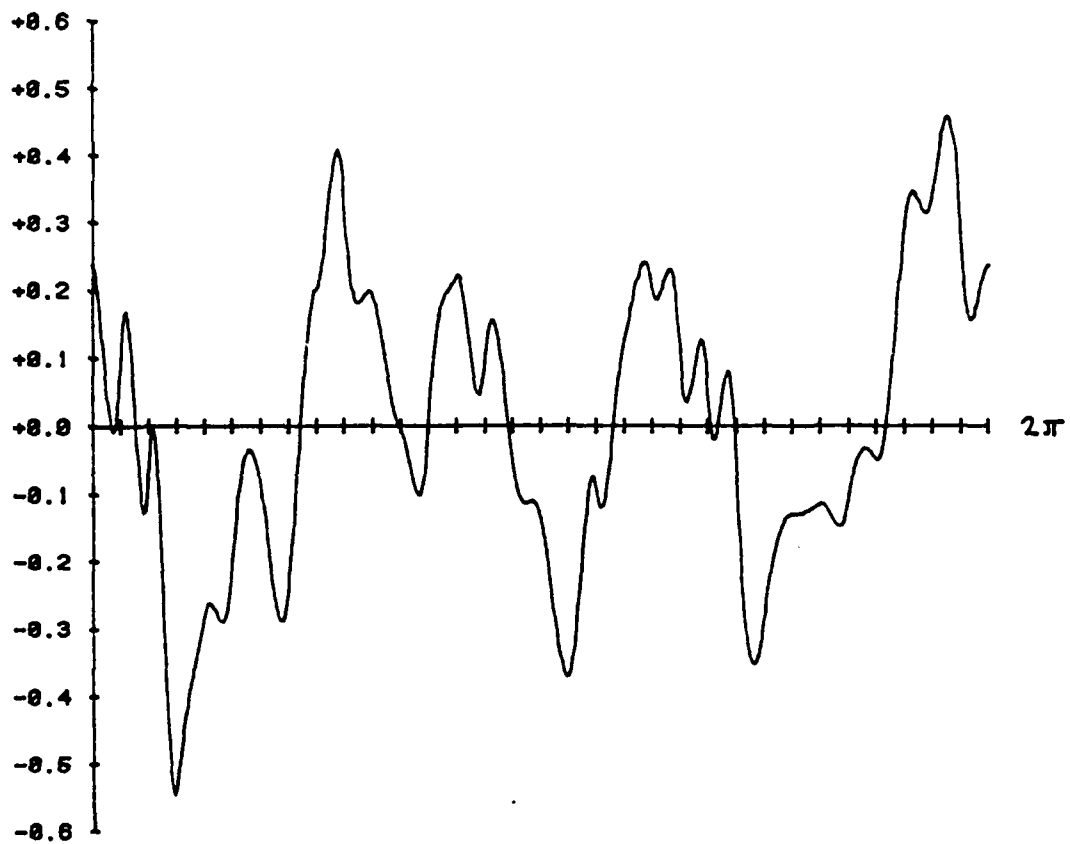


Figure 6.3(b). Calculated radial derivative V_r at $r = 1$ in experiment using comparatively large elements.

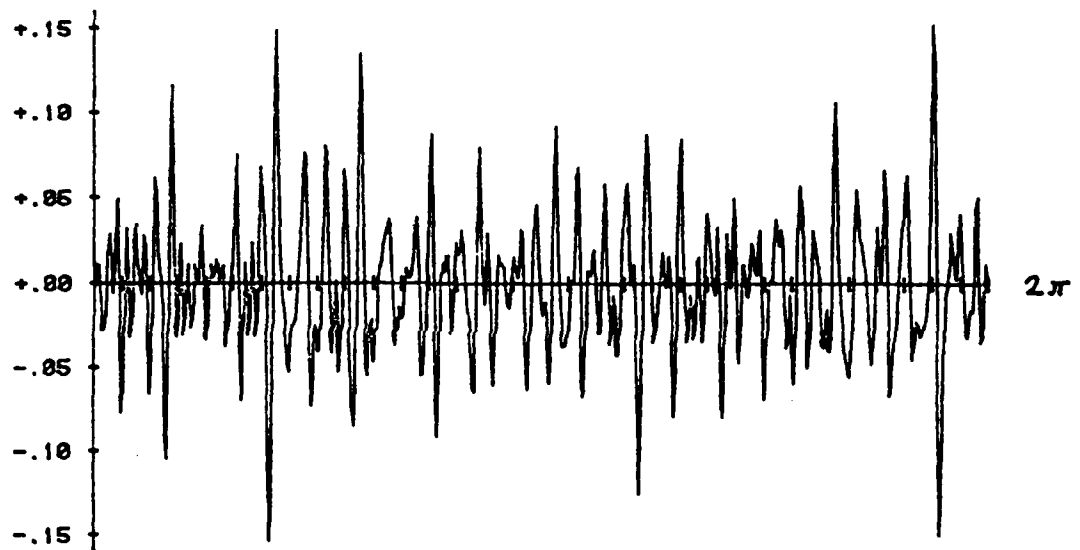


Figure 6.3(c). Difference calculated minus assumed radial derivative V_r at $r = 1$ in experiment using comparatively large elements.

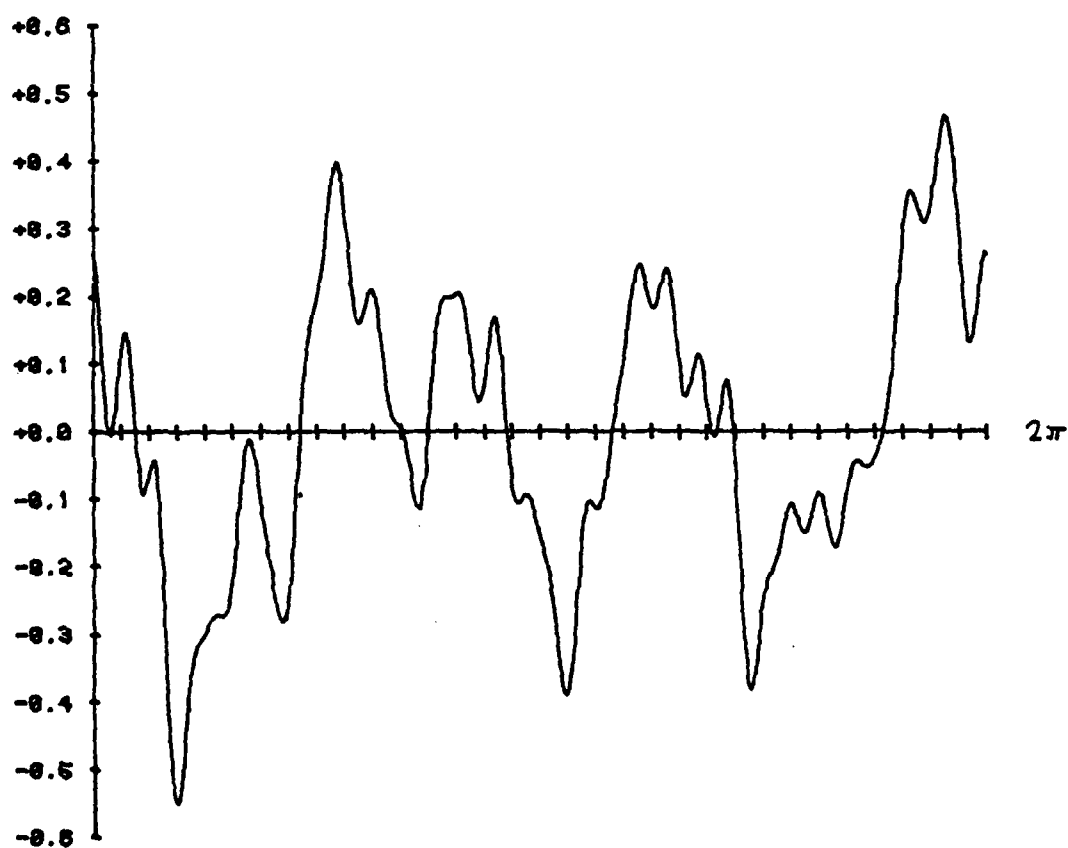


Figure 6.3(d). Radial derivative V_r of the assumed field at $r = 1$, truncated to circular harmonics of degree $n < 32$. Experiment using comparatively large elements.

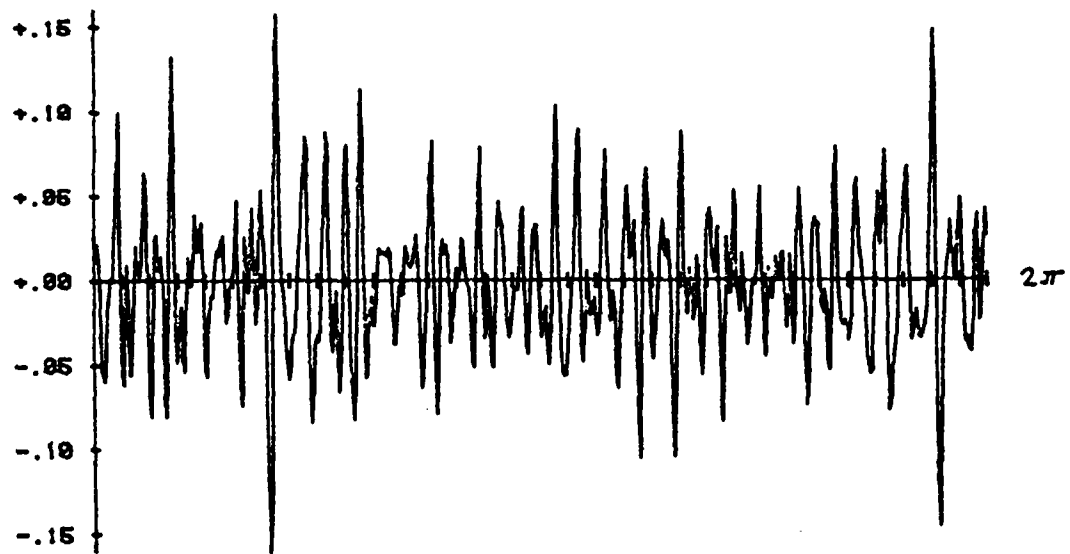


Figure 6.3(e). Contribution of circular harmonics of degree $n > 32$ to the radial derivative V_r of the assumed field at $r = 1$. Experiment using comparatively large elements.

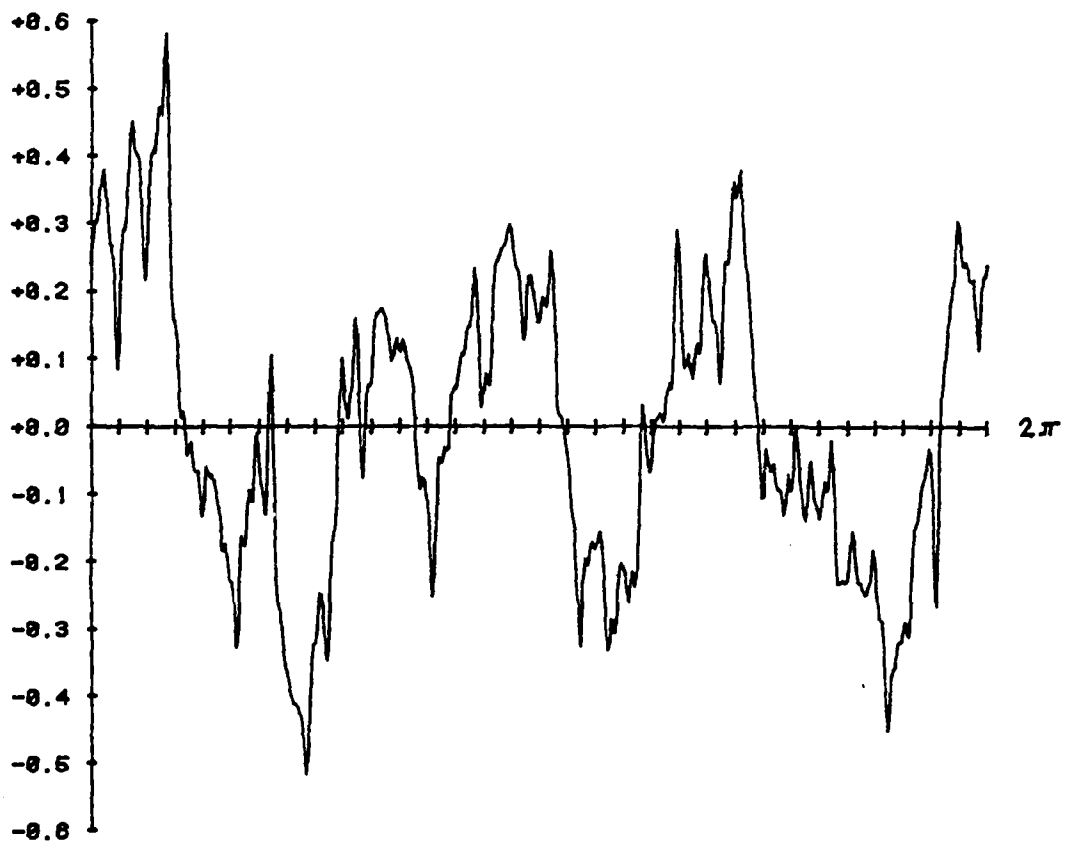


Figure 6.4(a). Assumed horizontal derivative V_y at $r = 1$ in experiment using comparatively large elements.

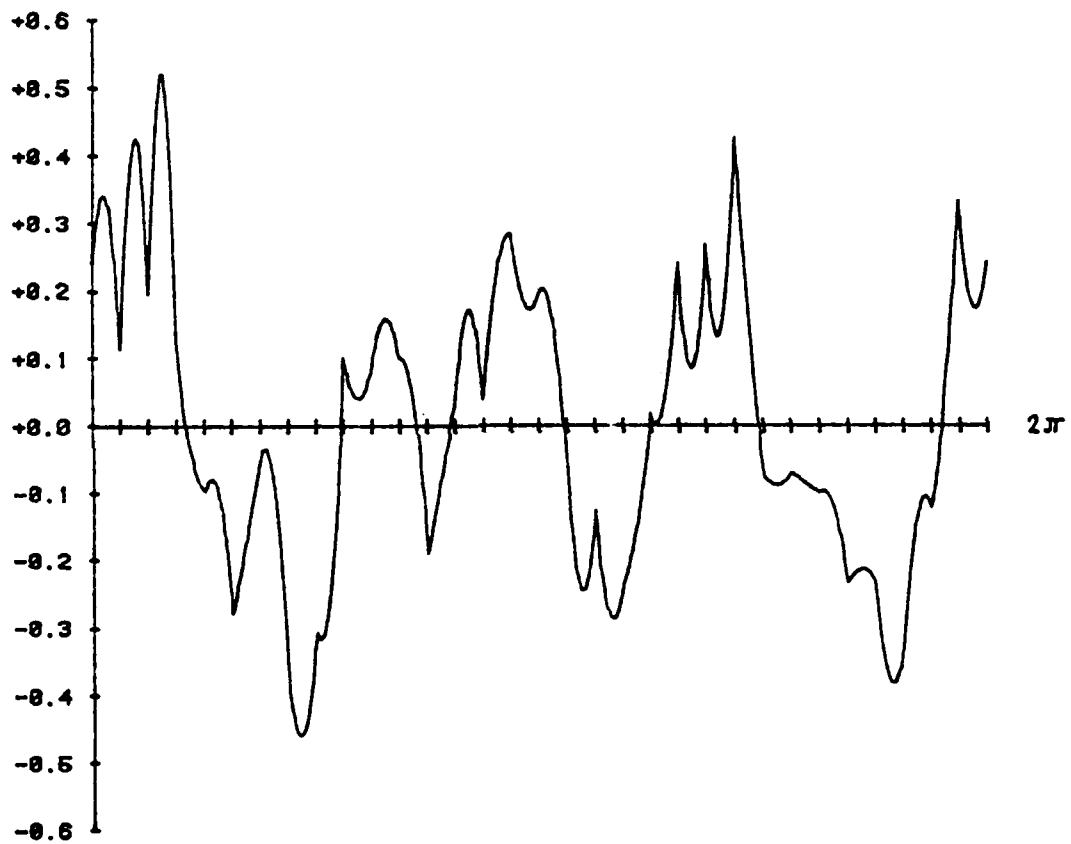


Figure 6.4(b). Calculated horizontal derivative V_y at $r = 1$ in experiment using comparatively large elements.

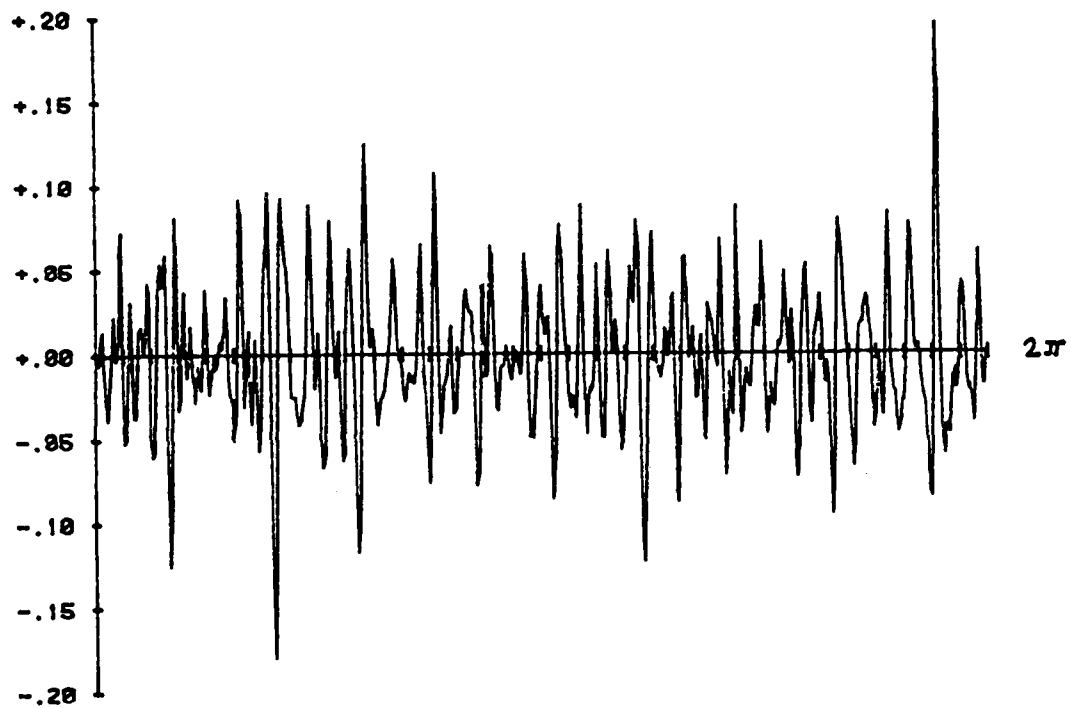


Figure 6.4(c). Difference calculated minus assumed horizontal derivative V_y at $r = 1$ in experiment using comparatively large elements.

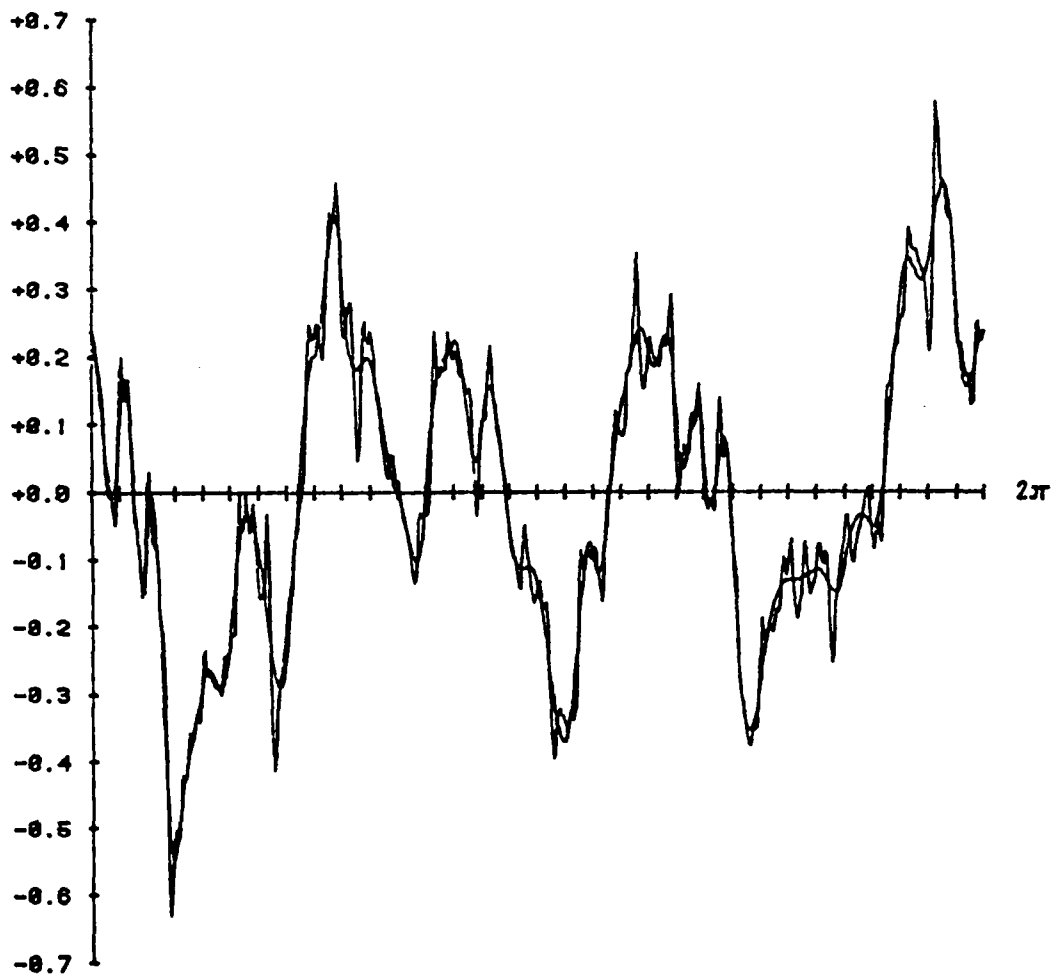


Figure 6.5. Superposition of assumed and calculated radial derivative V_r at $r = 1$ in experiment using comparatively large elements.

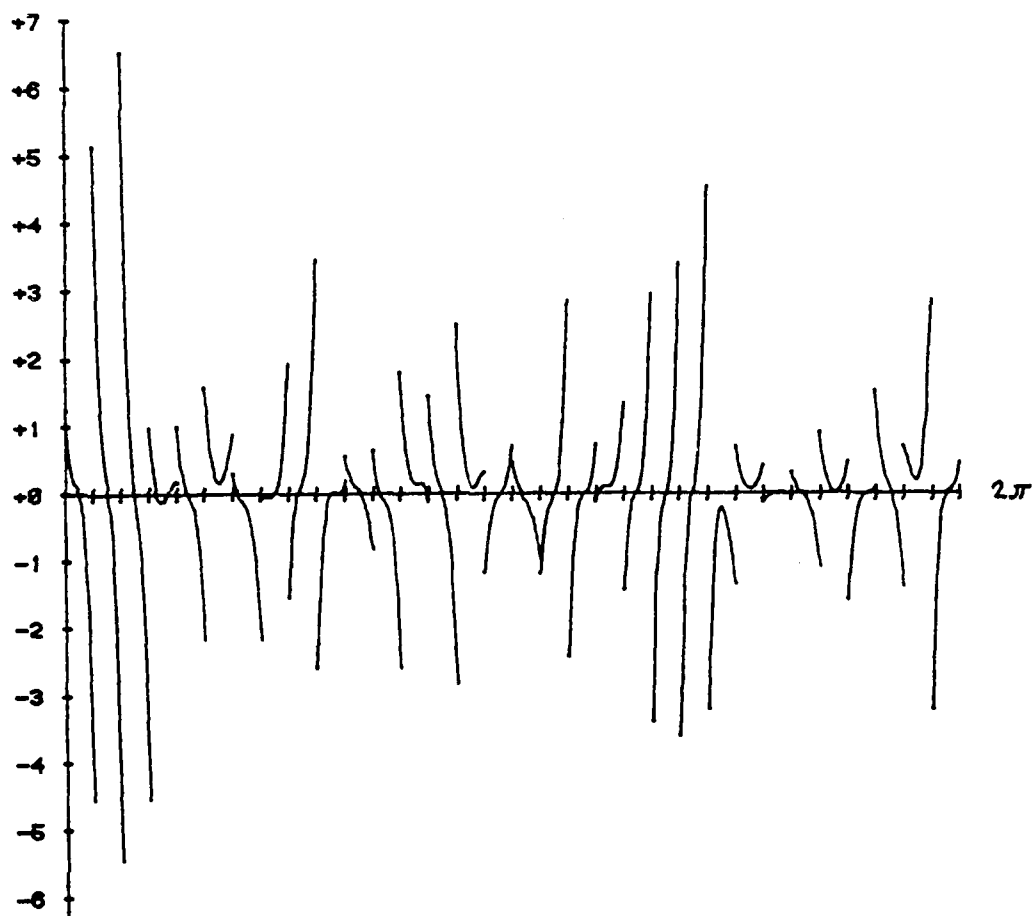


Figure 6.6. Calculated Laplacean ΔV at $r = 1$ in experiment using comparatively large elements.

In order to put the potential derived from finite elements to a further test, a number of orbits of passive mass-points was numerically integrated. A reference potential equal to

$$U(r) = \text{Log } \frac{1}{r} \quad (6.10)$$

was assumed. If $V_{\text{ASS}}(r, \varphi)$ and $V_{\text{CLC}}(r, \varphi)$ denote the assumed and calculated potential discussed previously, then the total potential used in the orbit calculations was either

$$W_{\text{ASS}}(r, \varphi) = U(r) + c_0 V_{\text{ASS}}(r, \varphi) \quad (6.11a)$$

or

$$W_{\text{CLC}}(r, \varphi) = U(r) + c_0 V_{\text{CLC}}(r, \varphi) \quad (6.11b)$$

By choosing

$$c_0 = 0.0003$$

an attempt was made to relate the reference potential and the disturbing potential in a way quantitatively similar to the real earth.

The equation of motion in polar coordinates is

$$\begin{aligned} \ddot{r} - r\dot{\varphi}^2 &= \frac{\partial W}{\partial r} \\ \ddot{\varphi} + 2\frac{\dot{r}}{r}\dot{\varphi} &= \frac{1}{r^2} \frac{\partial W}{\partial \varphi} \end{aligned} \quad (6.12)$$

For near circular orbits one may represent r and φ in the following way:

$$\begin{aligned} r &= r_0 + \Delta r \\ \varphi &= \dot{\varphi}_0 t + \Delta \varphi \end{aligned} \quad (6.13)$$

Thereby the two constants $r_0, \dot{\varphi}_0$ fulfill

$$r_0 \dot{\varphi}_0 = 1 \quad (6.14)$$

One proceeds to linearize the equation (6.12) obtaining:

$$\begin{aligned} \Delta \ddot{r} - \dot{\varphi}_0^2 \Delta r - 2r_0 \dot{\varphi}_0 \Delta \dot{\varphi} &= \frac{\Delta r}{r_0^2} + c_0 \frac{\partial V}{\partial r}(r_0 + \Delta r, \dot{\varphi}_0 t + \Delta \varphi) \\ \Delta \ddot{\varphi} + 2 \frac{\dot{\varphi}_0}{r_0} \Delta \dot{r} &= \frac{c_0}{(r_0 + \Delta r)^2} \frac{\partial V}{\partial \varphi}(r_0 + \Delta r, \dot{\varphi}_0 t + \Delta \varphi) \end{aligned} \quad (6.15)$$

We shall refer to (6.12) as the equation of the "exact orbit" whereas (6.15) are the equations of the "differential orbit".

Either version of the orbital equations was integrated by a high degree Runge-Kutta-type formula found in Henrici (1962), p. 171. When the calculated potential was used, and when the orbit crossed an element boundary, the step size was temporarily decreased by a factor of 1/4.

Figure 6.7(a) shows the results of an exact orbit calculation. The orbit passes through all 4 layers of the circular ring subdivided into elements of finite size. (Cf. figure 3.6(b)). The radial and angular deviations between the two solutions based on W_{CLC} and W_{ASS} are shown in figure 6.7(b). The deviation amounts to about 1 pp 300,000. This corresponds to a 20m-accuracy in the real world.

Figures 6.8(a)-(b) show in a similar way the results for a differential orbit with respect to $r_0 = 1.1, \dot{\varphi}_0 = 0.90909...$. It is a low orbit which passes about midway through the lowermost layer. We see that the accuracy is about 1 pp 1,500,000 or 4-5 m.

Remark: The deviations due to differing V_{CLC} and V_{ASS} of the orbits (exact as well as differential) are nearly linear in c_0 over a fairly wide range. Hence the assumption $c_0 = 0.0003$ can be changed by rescaling figures 6.7(b) and 6.8(b) appropriately.

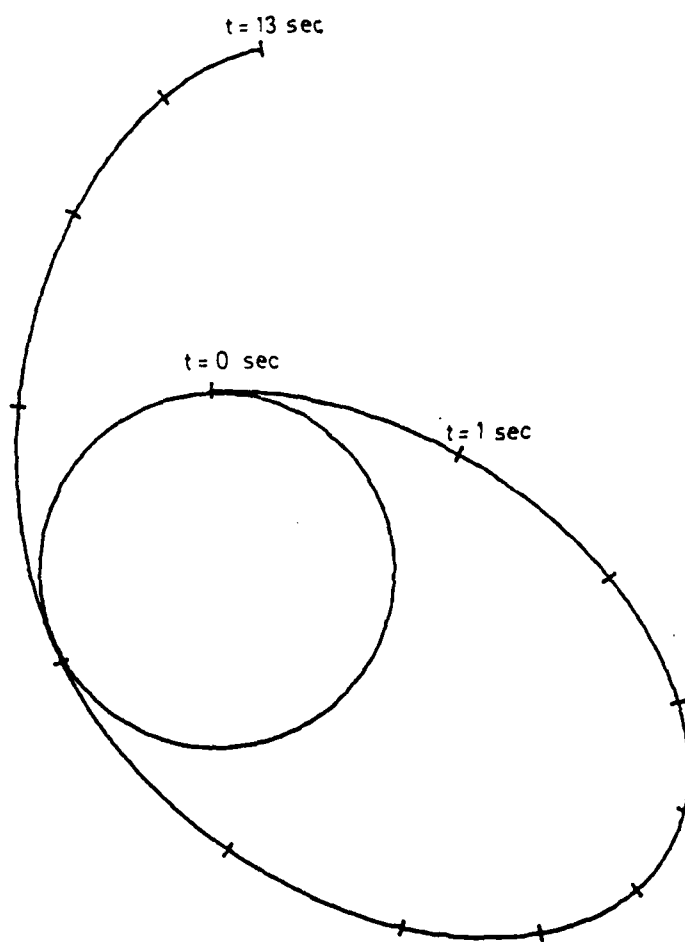


Figure 6.7(a). Exact orbit from experiment using comparatively large elements. Initial values are $r(0) = 1.01$, $\dot{r}(0) = 0$, $\varphi(0) = 0$, $\dot{\varphi}(0) = 1.55$. A step size of $\Delta t = 0.05$ seconds was used for the calculated potential and $\Delta t = 0.0125$ for the assumed potential. A value $c_0 = 0.0003$ was chosen. (Cf. equations (6.11)).

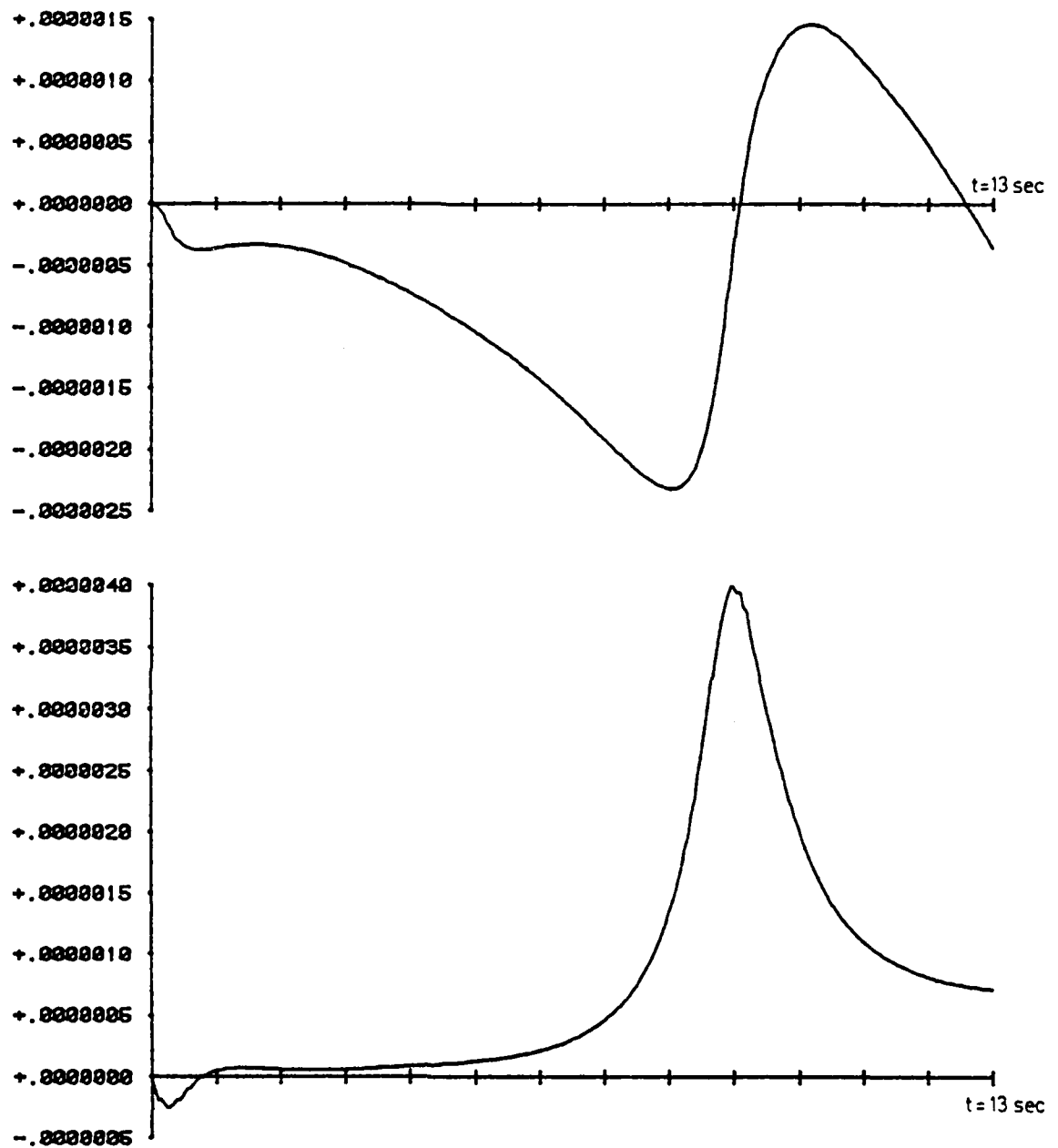


Figure 6.7(b). Difference in r (top) and ϕ (bottom) from calculated and assumed potential. Experiment using comparatively large elements.

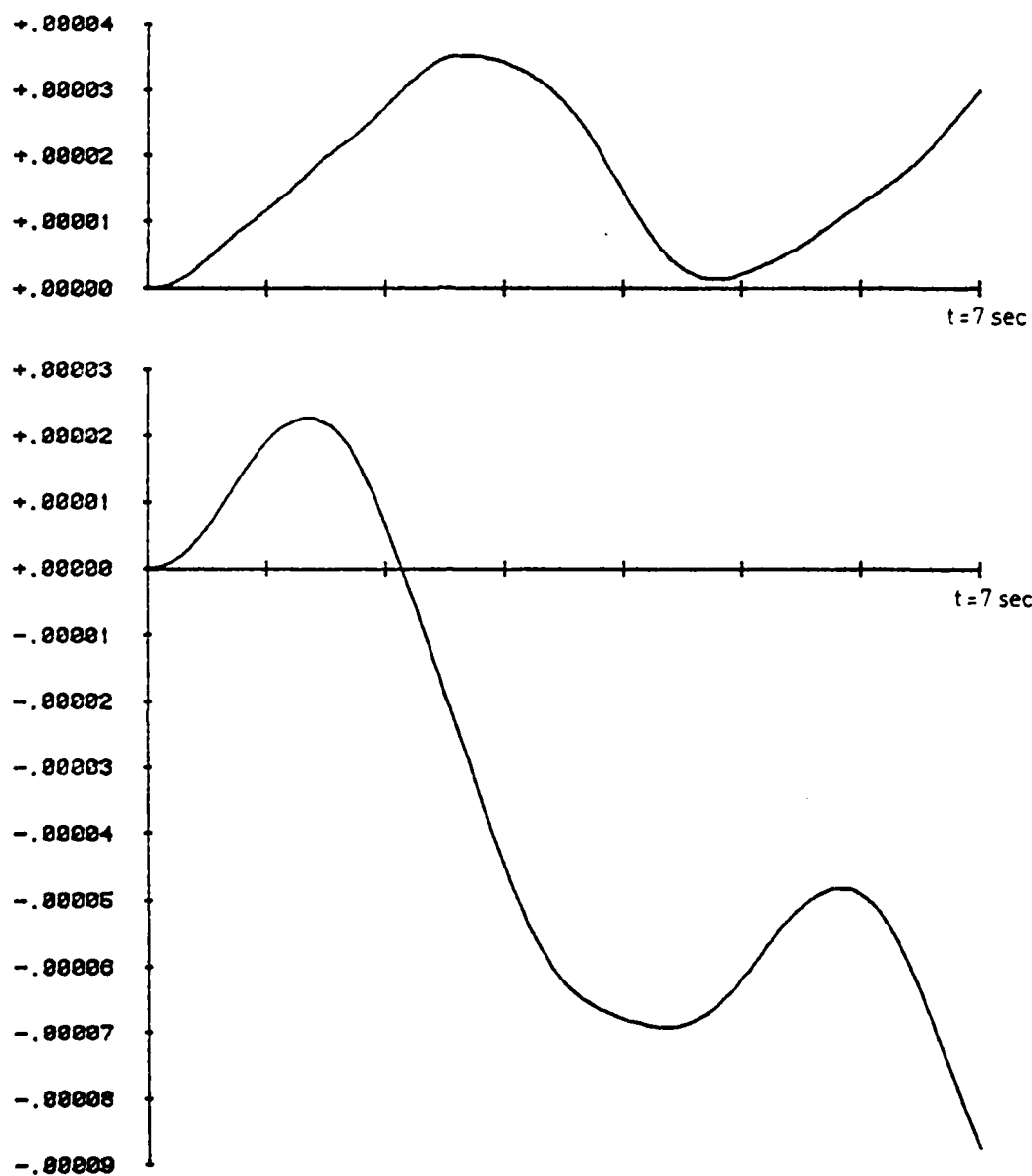


Figure 6.8(a). Radial increment Δr (top) and angular increment $\Delta \phi$ (bottom) obtained by integrating a differential orbit based on the assumed potential. Initial values: $r(0) = 1.1$, $\dot{r}(0) = 0$, $\phi(0) = 0$, $\dot{\phi}(0) = 0.90909\dots$. A step size of 0.05 sec was used for the calculated potential, and $\Delta t = 0.0125$ for the assumed potential. A value $c_0 = 0.0003$ was chosen.

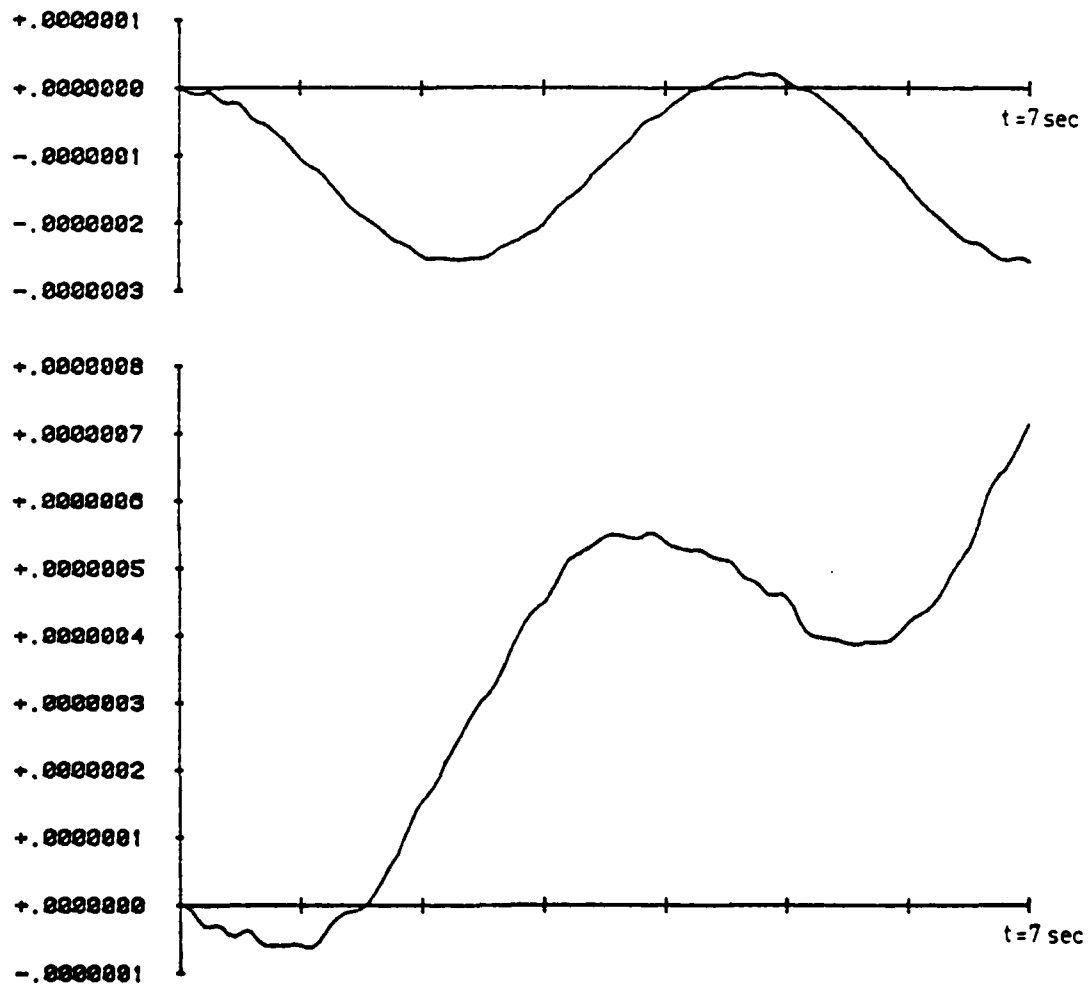


Figure 6.8(b). Difference in Δr (top) and Δg (bottom) of differential orbits based on calculated and assumed potential. Experiment using comparatively large elements.

Remark: Note that orbits in a central force field implied by (6.10) are not closed. In case of the differential orbit, the angle between successive pericenters (successive apocenters) is given by (cf. Arnold (1978), p. 37, problem 2)

$$\phi = \sqrt{2} \pi \quad (6.16)$$

Translated into time this gives

$$\frac{\phi}{\dot{\phi}_0} = 4.9 \text{ sec} \quad (6.17)$$

This is in agreement with figure 6.8(a).

We summarize some further results of this experiment in tables 6.2 and 6.3. Table 6.2 lists the magnitude of the nodal parameters and the magnitude of the deviations between calculated and assumed field. Maximal and rms quantities are listed per level. There are 7 levels of nodes. A layer of simple quads is bounded by two levels. A layer of compound quads has one additional level of nodes halfway between the two bounding levels. Level 1 corresponds to $r = 1$, level 7 to $r = r_{out} = 3.55884 \dots$

Level	V		V _r		V _y	
	max	rms	max	rms	max	rms
1	0.126944	0.066857	0.502961	0.198303	0.365901	0.201368
	0.002493	0.001005	0.098707	0.043960	0.092146	0.039320
2	0.078590	0.041004	0.178782	0.093193	0.212148	0.113458
	0.000258	0.000128	0.005493	0.002216	0.014307	0.005119
3	0.049725	0.025910	0.087696	0.046367	0.130977	0.067720
	0.000136	0.000055	0.000715	0.000291	0.002135	0.000799
4	0.031793	0.016990	0.045800	0.024350	0.075744	0.043491
	0.000084	0.000034	0.000101	0.000054	0.000380	0.000221
5	0.021928	0.011866	0.026655	0.014127	0.050927	0.029458
	0.000093	0.000033	0.000075	0.000038	0.000256	0.000126
6	0.009972	0.005744	0.008395	0.004658	0.020399	0.014113
	0.000045	0.000026	0.000037	0.000024	0.000612	0.000418
7	0.005620	0.003345	0.003601	0.002040	0.011373	0.007977
	0.000016	0.000010	0.000040	0.000022	0.000126	0.000090

Table 6.2

Maximal and rms values for assumed nodal parameters (top entry in each field) and for deviations from calculated values (bottom entries) at various levels of nodes. Experiment using comparatively large elements.

Table 6.3 compares some of the assumed potential coefficients a_n , b_n ; $n = 2, \dots, \text{NHARM} = 8$ with A_n , B_n obtained by the collocation procedure at $r = r_{\text{sur}}$.

n	A_n	B_n
	a_n Diff.	b_n Diff.
2	-0.016589	0.051241
	-0.016752	0.051570
	0.000163	-0.000329
3	-0.054197	0.041178
	-0.055401	0.042326
	0.001204	-0.001148
4	-0.000030	0.015842
	0.000313	0.019369
	-0.000343	-0.003527

(the remaining coefficients are not listed because the errors are comparable in size to the coefficients themselves)

Table 6.3
Comparison of assumed (top entries) and calculated (middle entries) harmonic coefficients. Experiment with comparatively large elements.

6.3.2. An experiment using comparatively small elements.

The input parameters were specified as follows

IDEGR	=	3	[same as before]
ISWINF	=	2	[same as before]
NLAYER	=	4	[same as before]
ITYPE(1)	=	1	[same as before]
ITYPE(2)	=	1	[same as before]
ITYPE(3)	=	2	[same as before]
ITYPE(4)	=	2	[same as before]
NELEM(1)	=	128	[increased by a factor of 4]
HFCT	=	1	[same as before]
NPOTC	=	64	[decreased by factor 1/2]
DAMP	=	1	[less damping than before]
NPINST	=	8	[same as before]
STKWGT	=	5	[same as before]
NHARM	=	32	[increased by a factor of 4]
NPINHA	=	8	[same as before]
HARWG0	=	50	[same as before]
HARWGI	=	5	[same as before]

Table 6.4 and figure 6.9 illustrate the geometry of the element partition. Note that r_{ovr} results in a smaller value than before.

LAYER	ITYPE	NELEM	RADIUS	DELPHI
1	1	128	1.00000	0.04909
2	1	128	1.04909	0.04909
3	2	64	1.10058	0.09817
4	2	32	1.20863	0.19635
$r_{avr} = 1.44595$				

Table 6.4. Element partition in experiment using comparatively small elements.

We now exhibit without much text the counterparts of figures 6.2-6.4, 6.7-6.8 and tables 6.2 and 6.3. These are the figures 6.10-6.14 and the tables 6.5-6.6.

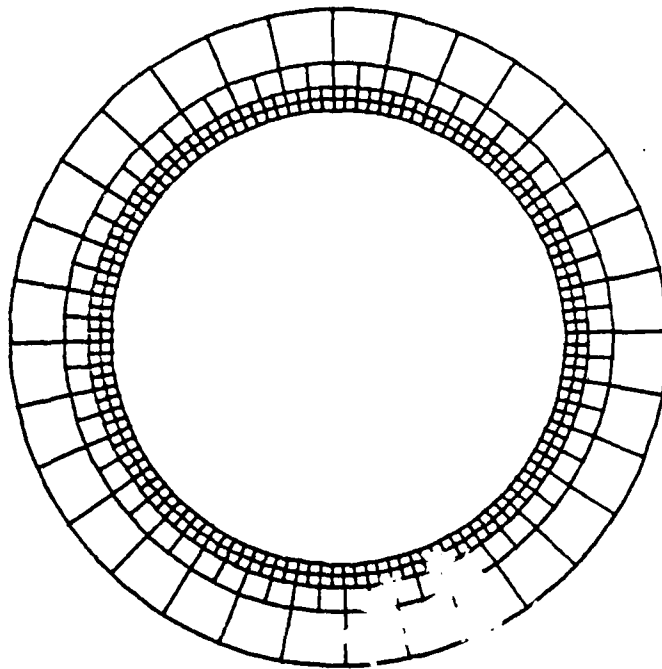


Figure 6.9. Element partition for experiment using comparatively small elements.

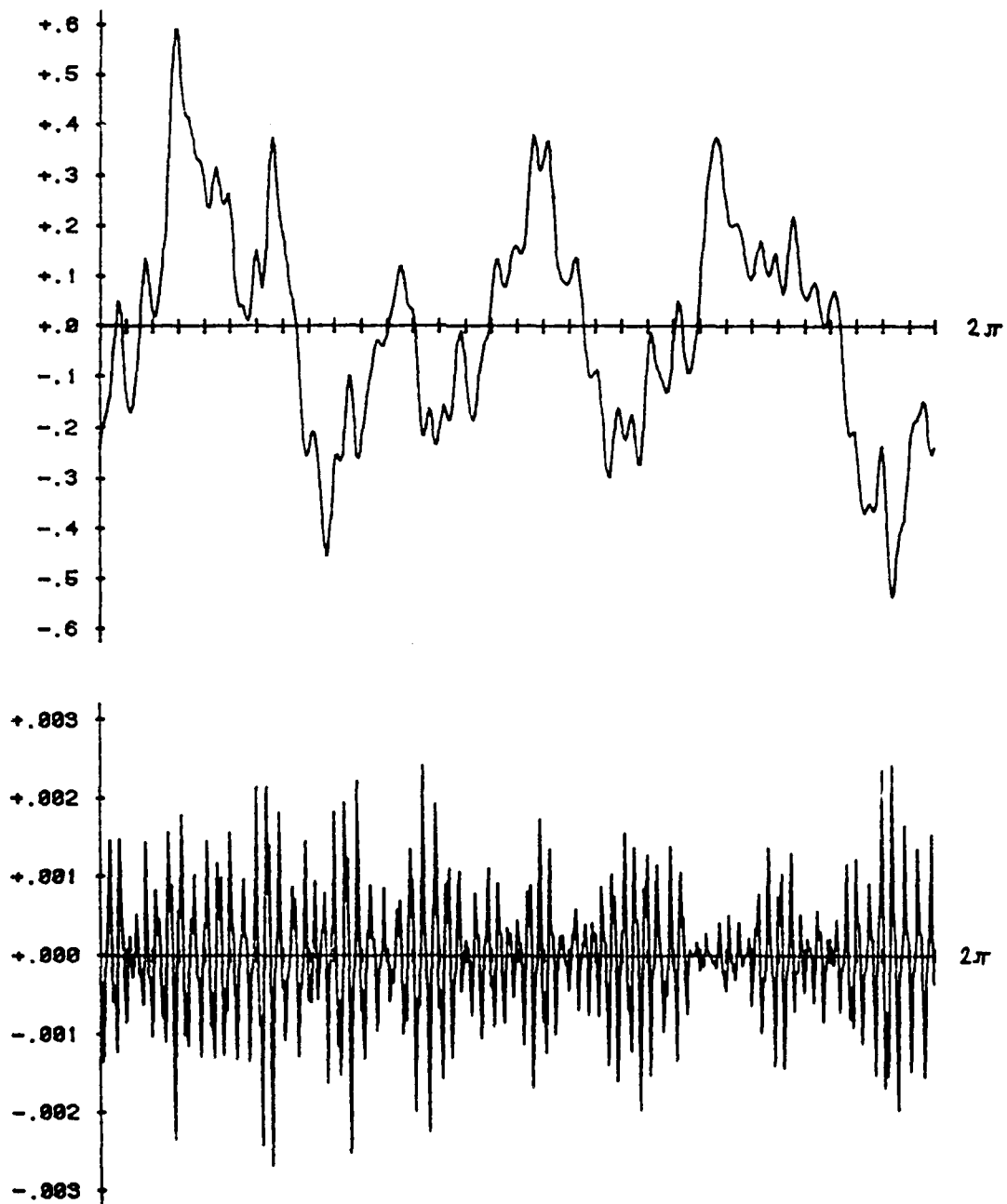


Figure 6.10. Assumed potential V (top) at $r = 1$ in experiment using comparatively small elements. The calculated V is graphically indistinguishable. The bottom figure shows the enlarged difference calculated minus assumed V .

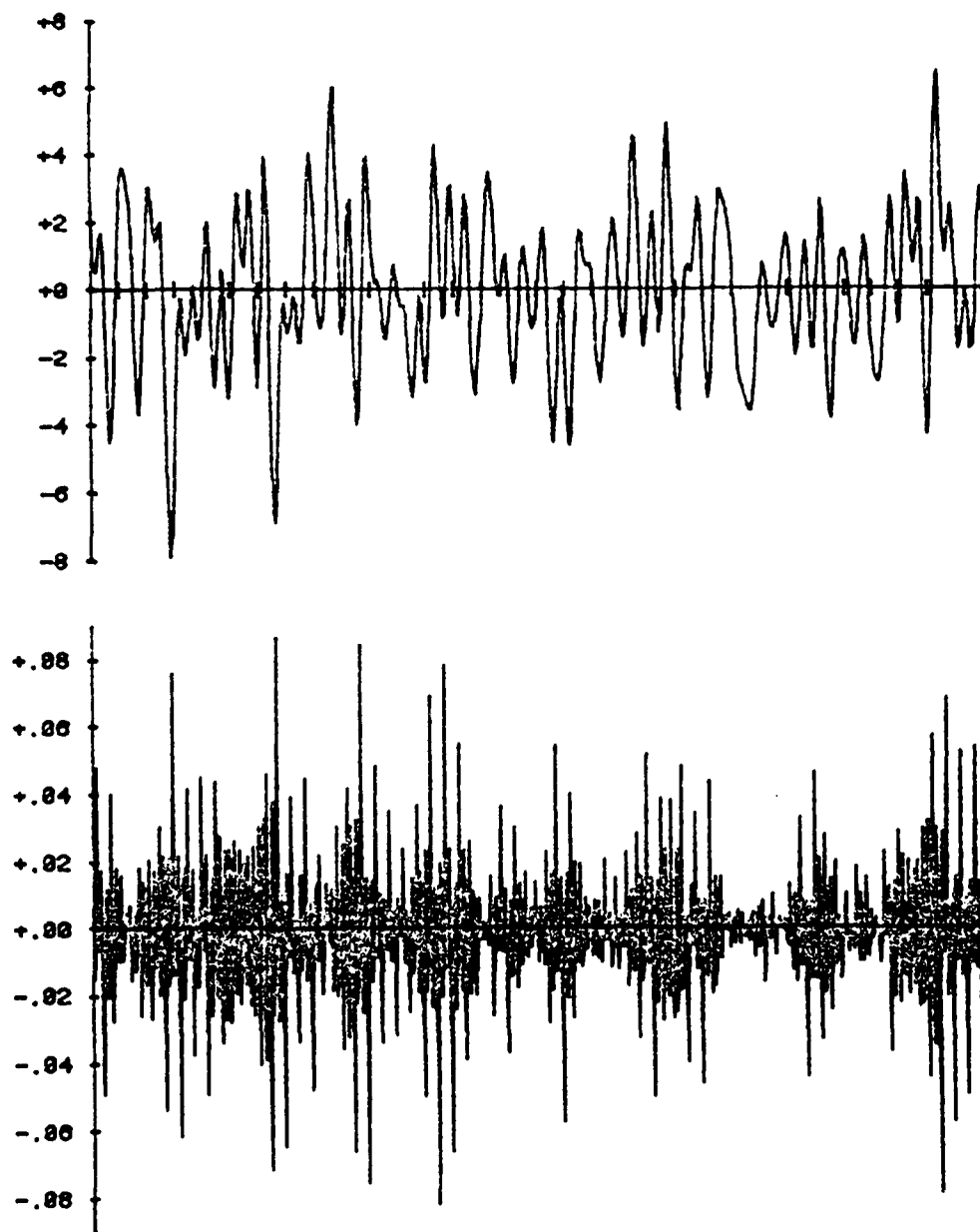


Figure 6.11. Assumed radial derivative V_r (top) at $r = 1$ in experiment using comparatively small elements. The calculated V_r is graphically nearly indistinguishable. The bottom figure shows the enlarged difference calculated minus assumed V_r at $r = 1$.

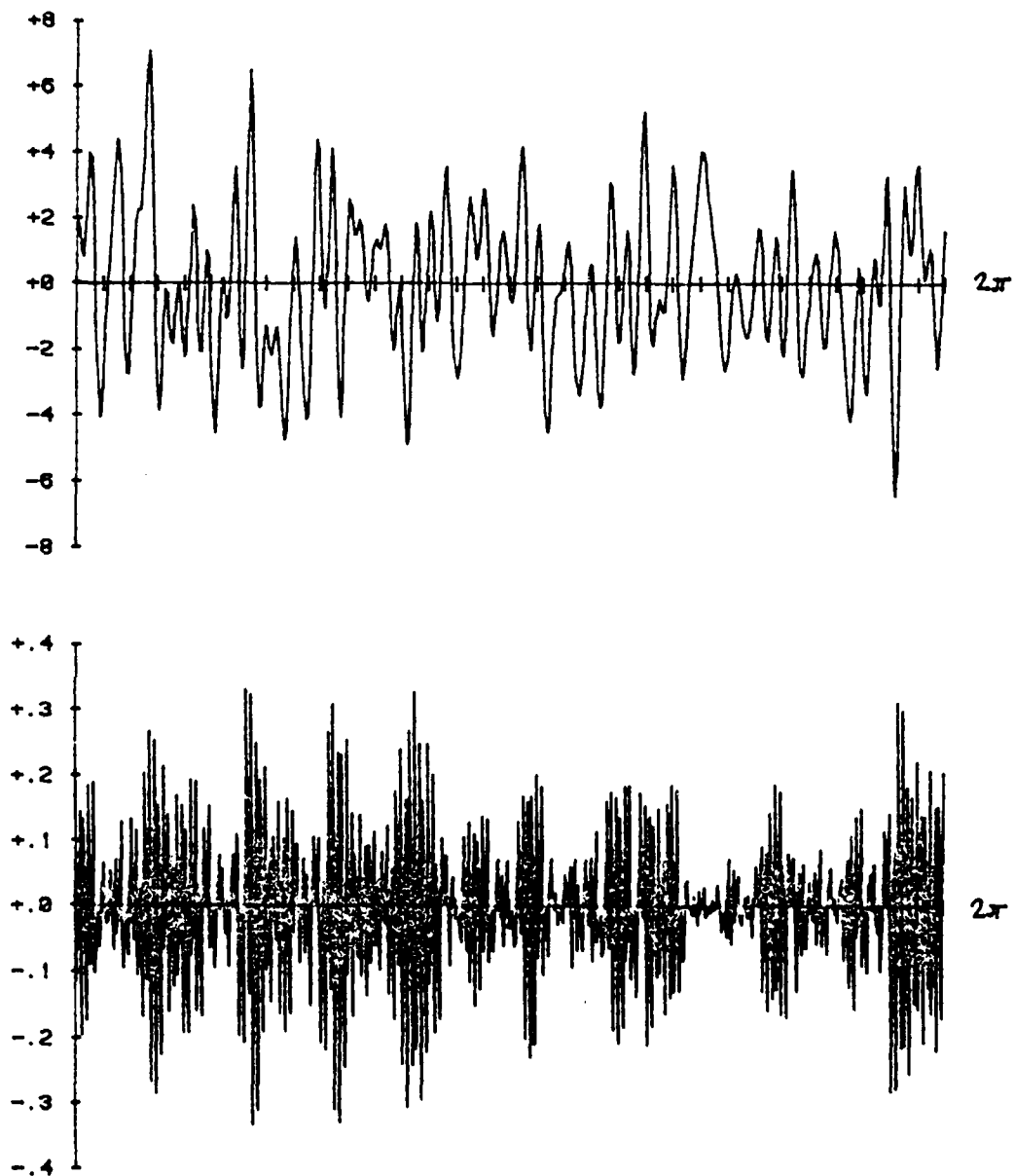


Figure 6.12. Assumed angular derivative V_θ (top) at $r = 1$ in experiment using comparatively small elements. The calculated V_θ is not shown. The bottom figure shows the enlarged difference calculated minus assumed V_θ at $r = 1$.

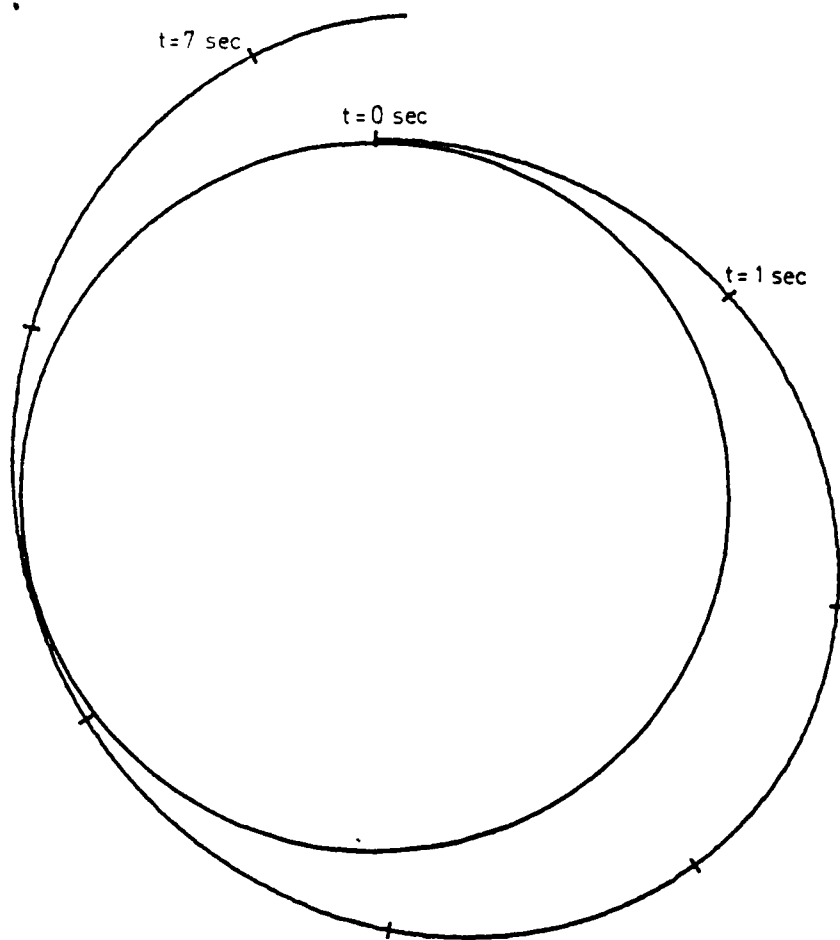


Figure 6.13(a). Exact orbit in experiment using comparatively small elements. Initial values are $r(0) = 1.01$, $\dot{r}(0) = 0$, $\varphi(0) = 0$, $\dot{\varphi}(0) = 1.15$. A step size of 0.0125 sec was used. A value $c_0 = 0.00003$ was chosen. (Cf. equations (6.11)).

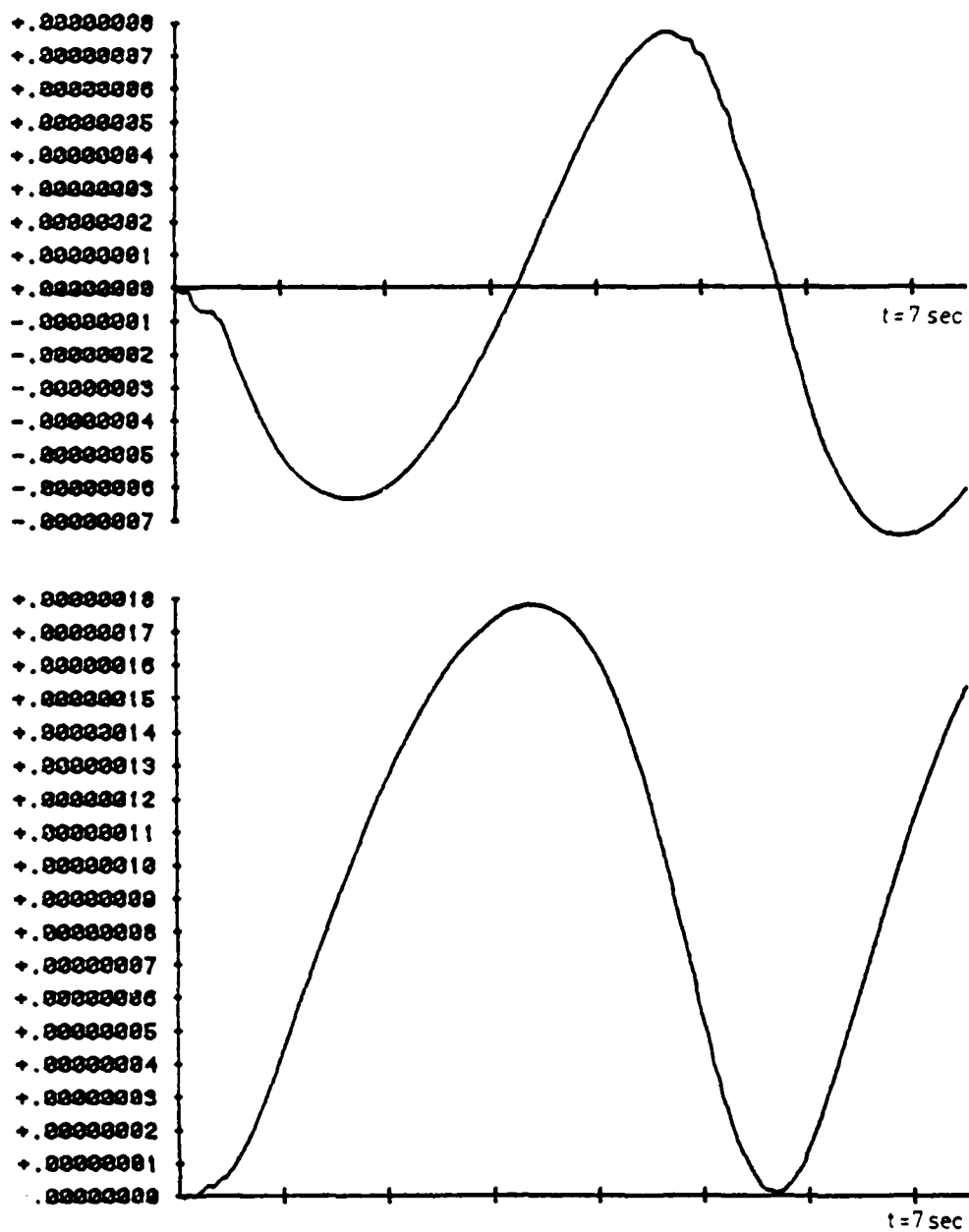


Figure 6.13(b). Difference in r (top) and φ (bottom) of exact orbits obtained from calculated and assumed potential. Experiment using comparatively small elements.

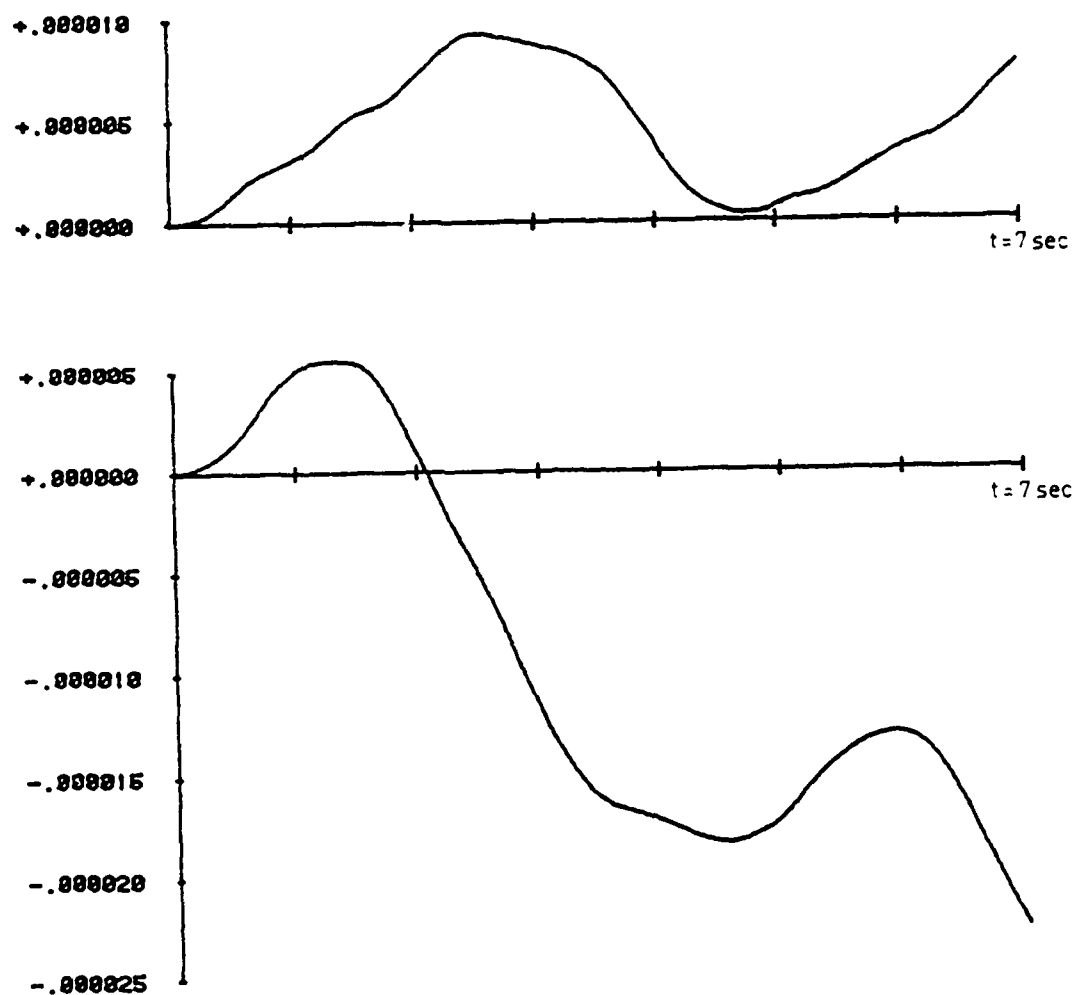


Figure 6.14(a). Radial increment Δr (top) and angular increment $\Delta \varphi$ (bottom) obtained by integrating a differential orbit based on the assumed potential. Initial values were $r(0) = 1.1$, $\dot{r}(0) = 0$, $\varphi(0) = 0$, $\dot{\varphi}(0) = 0.90909\dots$. A value $c_0 = 0.00003$ was chosen. Experiment using comparatively small elements.

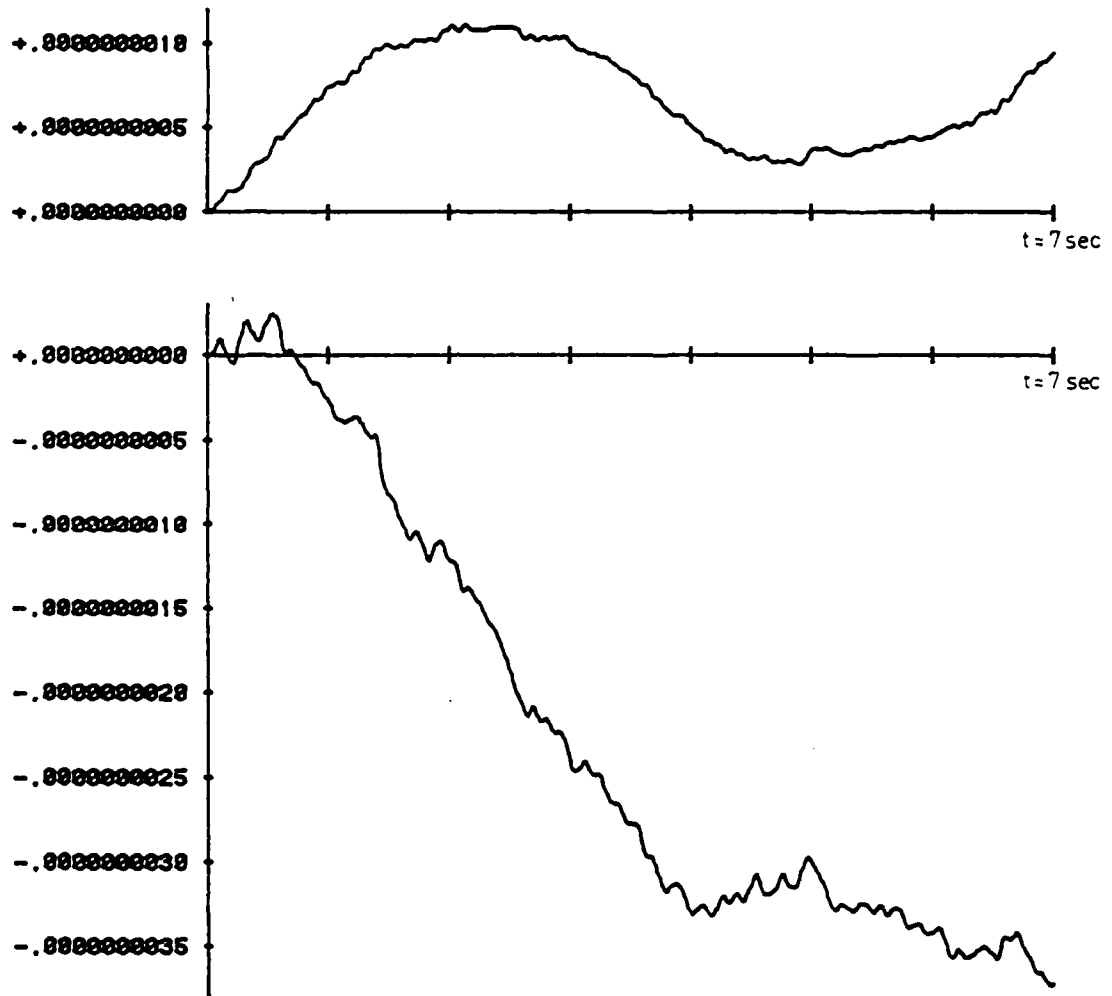


Figure 6.14(b). Difference in Δr (top) and $\Delta \phi$ (bottom) of differential orbits based on calculated and assumed potential. Experiment using comparatively small elements.

Note that the assumed potentials are different in the examples described in the previous and in this subsection. In subsection 6.3.1. we had NPOTC = 128 coefficients in the assumed potential and a damping factor of DAMP = 2. In this subsection NPOTC = 64 and DAMP = 1. The gradients of the present potential are about 10 times larger at $r = 1$ than in the earlier example. For this reason, the constant c_0 of equation (6.11) was chosen by a factor of 1/10 smaller than before. However, a remark given in section 6.3.1. carries over, implying that the figures 6.13(b) and 6.14(b) may be rescaled in proportion to any change of c_0 .

Due to the differences in the assumed potentials the differences in the results are not only due to the change of the element partition. Hence the results are not immediately compared. In the next section a number of experiments will be described summarily. From this additional information it may be concluded, that the transition from NELEM(1) = 32 to a value of 128 improves the results by about 1 digit at $r = 1$.

Figures 6.13 and 6.14 showing the results of the orbit calculations demonstrate an accuracy of about 1 pp 5,000,000 in case of the exact orbit. This corresponds to 1-2 m. The differential orbit is good to 1 pp 300,000,000. This correspond to 2 cm.

Level	V		V _r		V _g	
	max	rms	max	rms	max	rms
1	0.560348	0.214973	5.990022	2.260387	6.250663	2.255706
	0.002427	0.000882	0.046030	0.016975	0.328472	0.153894
2	0.384302	0.171249	2.389440	0.797211	2.324042	0.834182
	0.000698	0.000250	0.053510	0.019825	0.011507	0.004792
3	0.297927	0.143964	1.330956	0.502973	1.334445	0.553725
	0.000094	0.000026	0.004152	0.001534	0.005652	0.001970
4	0.243919	0.123493	0.853614	0.368318	0.912405	0.427124
	0.000053	0.000023	0.001245	0.000492	0.005634	0.001950
5	0.209708	0.107151	0.606866	0.286809	0.696430	0.348634
	0.000044	0.000019	0.000588	0.000227	0.001248	0.000473
6	0.156166	0.082076	0.348728	0.184487	0.458199	0.248361
	0.000091	0.000046	0.000463	0.000222	0.003435	0.001451
7	0.122152	0.064510	0.240170	0.127450	0.340255	0.187136
	0.000058	0.000031	0.000102	0.000052	0.000789	0.000376

Table 6.5

Maximal and rms values for assumed nodal parameters (top entry in each field) and for deviations from calculated values (bottom entries) at various levels of nodes. Experiment using comparatively small elements.

n	A _n a _n Diff.	B _n b _n Diff.	n	A _n a _n Diff.	B _n b _n Diff.
2	-0.033504 -0.033505 0.000001	0.103139 0.103140 -0.000001	13	-0.005873 -0.006296 0.000423	-0.026646 -0.027436 0.000790
3	-0.166199 -0.166203 0.000004	0.126972 0.126977 -0.000005	14	0.009536 0.010171 -0.000635	0.027354 0.027769 -0.000414
4	0.001247 0.001250 -0.000003	0.077468 0.077477 -0.000009	15	-0.001018 -0.000838 -0.000180	0.025260 0.025864 -0.000604
5	-0.035952 -0.035958 0.000006	-0.058543 -0.058558 0.000015	16	-0.013724 -0.013725 0.000001	-0.010341 -0.011124 0.000783
6	0.023489 0.023507 -0.000017	0.072583 0.072625 -0.000042	17	-0.007358 -0.007605 0.000247	-0.008160 -0.007283 -0.000877
7	-0.025322 -0.025337 0.000015	-0.013856 -0.013859 0.000003	18	-0.019404 -0.020553 0.001149	-0.015308 -0.014436 0.000872
8	0.031356 0.031431 -0.000075	-0.051497 -0.051597 0.000100	19	0.022968 0.023800 -0.000832	-0.018957 -0.020617 0.001660
9	0.052027 0.052222 -0.000195	-0.032417 -0.032523 0.000106	20	-0.012259 -0.012340 0.000081	-0.015923 -0.016005 0.000082
10	0.003373 0.003460 -0.000087	0.017371 0.017520 -0.000149	21	-0.001372 -0.001379 0.000006	0.007838 0.007600 0.000238
11	0.012190 0.012312 -0.000122	0.018735 0.018955 -0.000220	22	-0.020856 -0.019737 -0.001119	0.013256 0.011703 0.001553
12	0.000715 0.000830 -0.000016	-0.001960 -0.002129 0.000169	23	-0.004177 0.000844 -0.005021	0.005744 0.008234 -0.002490

Table 6.6
Comparison of assumed (top entries) and calculated
(middle entries) harmonic coefficients. Experiment
using comparatively small elements.

6.4. Summary of other experiments.

6.4.1. Experiments using bicubics.

Table 6.7 lists input parameters and resulting accuracies for a number of experiments using bicubics.

Experiment	1	2	3	4	5	6	7	8	9
ISWINF	1	1	1	2	1	2	1	1	1
NLAYER	4	4	4	4	4	4	5	4	4
ITYPE(1)	1	1	1	1	1	1	1	1	1
ITYPE(2)	1	1	1	1	1	1	1	1	1
ITYPE(3)	2	2	2	2	2	2	2	2	2
ITYPE(4)	2	2	2	2	2	2	2	2	2
ITYPE(5)	/	/	/	/	/	/	2	/	/
NELEM(1)	32	32	32	32	32	64	64	128	128
HFCT	1	1	1	1	1	1	0.75	1	1
ϵ_{our}	3.56	3.56	3.56	3.56	3.56	2.01	2.72	1.45	1.45
NPOTC	16	32	128	128	128	32	64	64	128
DAMP	1	1	2	2	1	1	1	1	1
NPINST	8	8	8	8	8	8	8	8	8
STKWGT	5	5	5	5	5	25	5	5	5
NHARM	8	8	8	/	8	/	8	32	32
NPINHA	8	8	8	/	8	/	8	8	8
HARWG0	50	50	50	/	50	/	50	50	50
HARWG1	15	15	5	/	5	/	5	5	5

(table continued on next page)

Experiment		1	2	3	4	5	6	7	8	9	
A, B	2	1E-3	2E-3	3E-4	/	2E-3	/	8E-4	1E-6	1E-5	
	3	4E-3	5E-3	1E-3	/	6E-3	/	2E-3	5E-6	6E-6	
	4	1E-2	2E-2	4E-3	/	2E-2	/	1E-2	9E-6	8E-6	
accuracy at levels	V	1	1E-3	2E-2	1E-3	1E-3	4E-2	1E-3	1E-2	9E-4	7E-3
		2	5E-4	3E-3	1E-4	1E-4	3E-3	4E-4	8E-4	2E-4	1E-3
		3	1E-4	1E-3	5E-5	6E-5	1E-3	5E-5	5E-4	2E-5	3E-4
		4	8E-5	7E-4	3E-5	4E-5	7E-4	4E-5	4E-4	2E-5	1E-4
		5	9E-5	5E-4	3E-5	4E-5	5E-4	4E-5	3E-4	2E-5	5E-5
		6	5E-5	4E-4	2E-5	1E-5	4E-4	7E-5	2E-4	5E-5	5E-5
		7	9E-6	2E-4	9E-6	1E-5	3E-4	7E-5	2E-4	3E-5	3E-5
		8	/	/	/	/	/	/	2E-4	/	/
		9	/	/	/	/	/	/	9E-5	/	/
	V _r	1	7E-3	5E-1	4E-2	4E-2	3	1E-2	7E-1	2E-2	1
		2	8E-3	5E-2	2E-3	2E-3	5E-2	1E-2	2E-2	2E-2	1E-1
		3	6E-4	6E-3	3E-4	3E-4	7E-3	1E-3	5E-3	2E-3	1E-2
		4	2E-4	1E-3	5E-5	5E-5	1E-3	3E-4	2E-3	5E-4	2E-3
		5	1E-4	5E-4	4E-5	4E-5	5E-4	2E-4	1E-3	2E-4	7E-4
		6	6E-5	2E-4	2E-5	3E-5	3E-4	9E-5	7E-4	2E-4	2E-4
		7	5E-5	2E-4	2E-5	1E-5	2E-4	2E-4	5E-4	5E-5	7E-5
		8	/	/	/	/	/	/	3E-4	/	/
		9	/	/	/	/	/	/	2E-4	/	/
	V _g	1	6E-2	8E-1	4E-2	4E-2	2	1E-1	9E-1	2E-1	1
		2	9E-4	1E-1	5E-3	5E-3	2E-1	3E-3	6E-2	5E-3	2E-1
		3	2E-3	2E-2	8E-4	8E-4	2E-2	2E-3	1E-2	2E-3	2E-3
		4	1E-3	2E-3	2E-4	2E-4	3E-3	1E-3	6E-3	2E-3	3E-3
		5	5E-4	1E-3	1E-4	1E-4	1E-3	3E-4	2E-3	5E-4	8E-4
		6	1E-3	1E-3	4E-4	3E-4	2E-3	1E-3	3E-3	2E-3	2E-3
		7	3E-4	6E-4	8E-5	2E-4	7E-4	3E-4	6E-4	4E-4	4E-4
		8	/	/	/	/	/	/	2E-3	/	/
		9	/	/	/	/	/	/	9E-4	/	/

Table 6.7
Summary of experiments using bicubics. Experiment No. 3 and No. 8 are those discussed in detail in sections 6.3.1. and 6.3.2.

6.4.2. Experiments using biquintics.

Table 6.8 lists input parameters and resulting accuracies for a number of experiments using biquintics. The parameters NLAYER (=4), ITYPE(1) (=1), ITYPE(2) (=1), ITYPE(3) (=2), ITYPE(4) (=2) are not repeatedly listed because the indicated values were the same for all experiments documented in this table.

Experiment	1	2	3	4	5
NELEM(1)	32	32	32	32	32
HFCT	1	0.75	0.75	1	1
r_{out}	3.56	2.71	2.71	3.56	3.56
NPOTC	32	128	128	128	128
DAMP	1	2	1	1	1
NPINST	8	16	16	16	8
STKWGT	1	1	1	1	1
NHARM	8	8	8	8	8
NPINHA	8	8	8	8	8
HARWG0	10	10	10	10	10
HARWG1	2.5	2.5	2.5	2.5	2.5
N89	9	9	9	9	8
FLPWGL	1E-3	1E-3	1E-3	1E-3	1E-3
FLPWGU	3E-3	1E-4	1E-4	1E-4	1E-4

(table continued on next page)

Experiment		1	2	3	4	5	
A, B	2	1E-3	1E-5	4E-4	5E-4	3E-2	
	3	2E-4	4E-6	2E-4	3E-4	2E-3	
	4	3E-4	3E-5	4E-4	6E-4	5E-3	
accuracy at levels	V	1	2E-3	4E-4	3E-2	3E-2	3E-2
		2	2E-3	7E-6	2E-4	5E-4	2E-3
		3	1E-3	7E-6	2E-4	3E-4	1E-3
		4	9E-4	5E-6	2E-4	2E-4	7E-4
		5	6E-4	4E-6	1E-4	1E-4	5E-4
		6	2E-4	2E-6	6E-5	5E-5	2E-4
		7	9E-5	1E-6	4E-5	3E-5	2E-3
	V _r	1	1E-2	3E-2	2	2	2
		2	9E-3	3E-4	1E-2	3E-2	2E-2
		3	3E-3	4E-5	2E-3	5E-3	6E-3
		4	1E-3	8E-6	3E-4	2E-4	9E-4
		5	9E-4	6E-6	2E-4	2E-4	6E-4
		6	3E-4	6E-6	8E-5	6E-5	2E-4
		7	6E-5	2E-6	3E-5	6E-5	1E-4
	V _φ	1	4E-2	3E-2	2	2	2
		2	1E-3	1E-4	4E-3	7E-3	9E-3
		3	2E-3	2E-5	9E-4	1E-3	2E-3
		4	1E-3	5E-5	5E-4	3E-4	1E-3
		5	8E-4	2E-5	3E-4	2E-4	9E-4
		6	3E-4	6E-5	2E-4	1E-4	4E-4
		7	2E-4	2E-5	1E-4	2E-5	3E-3

Table 6.8. Summary of experiments using biquintics.
 Additional input parameters were NLAYER = 4,
 ITYPE(1) = 1, ITYPE(2) = 1, ITYPE(3) = 2, ITYPE(4) = 2.

6.4.3. Discussion.

It is interesting to compare experiments based on the same assumed field, e.g. the bicubic-experiments 5 and 9. (The field at $r = 1$ is shown in figures 6.10-6.12. Also table 6.5 applies to this field). It is seen that the transition from 32 elements in the lowermost layer to 128 elements results in an increase of accuracy by one digit. On the other hand, experiments 8 and 9 differ only with respect to NPOTC. One recognizes that experiment 9 loses accuracy in the lower layers due to a rougher field there. However the accuracy in the upper layers is about the same. This confirms our reasoning that the finite element solution regularizes the field without affecting the lower frequencies. A comparison of experiments 3 and 4 shows that the representation of the field in the outer zone by spherical harmonics is about equivalent to representation by elements extending to infinity. However it must be borne in mind that the computational effort is less in the case of elements extending to infinity.

Experiment No. 5 using bicubics and experiment No. 4 using biquintics are otherwise based mostly on equal parameters. A comparison shows that the improvement coming from biquintics is marginal.

Experiments No. 4 and No. 5 of the biquintics-table show that fixing the Laplacean at zero gave worse results. It is particularly interesting how poorly the harmonic coefficients were recovered in experiment 5. In all experiments using biquintics the recoverage of the harmonic coefficients was inferior to that in the experiments using bicubics. The reason is not yet completely clarified.

6.4.4. A word on the Computer programs.

The programs for the Amdahl 470 V/6-II were written in an extended FORTRAN. They comprised about 6000 statements. Double precision was used throughout. Coding, testing and processing required about 6 weeks. No attempt was made to optimize speed by using assembly code in the inner loop of the equation solver. The CPU time for the largest experiment was 83 seconds. It required the solution of a linear system with 2368 unknowns. Recherches show that about half of the time was spent on evaluating the assumed potential and its derivatives for comparison purposes. The time spent on the solution of the linear system must have been less than 40 seconds. A full system of this size would have required a CPU of more than 74 minutes. Hence a factor of about 1/100 has been gained by the nested dissection method.

The programs for the post-analysis on the WANG 2200 VP-desk-top computer were written in an extended BASIC, called WANG BASIC-2. About 500k bytes of code were assembled requiring an effort of about 3 weeks. Doing the post-analysis on a large computer would have required a time span of probably 3 months. Small desk top computers have powerful editing facilities, they offer instant response during editing, and instant diagnosis during testing. It is a nonsense to solve small problems on large computers.

6.4.5. Further desirable experiments.

Time pressure did not allow to conduct all experiments the author had originally in mind. It was intended to generate "assumed potentials" other than those defined by circular harmonics. For example it would be interesting to see how the method performs for a potential generated by

buried dipoles. Also potentials with other types of near singularities could be tried. Another test could have been devoted to irregularly distributed measurement locations, to noisy data, and to less regular element partitions. Comparisons with the results of other methods applied to the same 2-dimensional problem are missing. Finally, tests in 3 dimensions should be performed. Unfortunately there is a limit to the amount of work a single person can do in a year, in addition to teaching and administrating. It is intended to continue the experiments with the help of my coworkers and students and to give a more complete report at a later occasion.

7. A proposed hybrid method.

7.1. Revision of the surface layer method.

There is an obvious way to introduce finite elements to the surface layer method, namely to represent the surface layer density in terms of 2-dimensional finite elements. Bilinear functions defined over rectangles in the φ, λ plane would ensure a continuous surface layer density. Bicubics would give a C^1 function. No improvement in computational efficiency can be expected from such a procedure. The normal equations would still be fully occupied. The number of parameters per node increases from one to four. Hence block areas can be chosen 4 times as large as in the case of constant densities within blocks. The benefit would be a field having continuous derivatives down to the surface of computation.

7.2. Multipole layer.

Keeping in mind that the potential V to be represented is actually a disturbing potential, and that frequently the reference potential U is chosen in a way that V is free of harmonics up to a certain degree $N-1$, it is tempting to try a surface layer density such that the generated potential is free of harmonics up to degree $N-1$. Such densities are available, though not in the form of a single layer. Layers of multipoles must be chosen. The potential generated by a dipole is

$$V = \frac{\mu \cos \varphi}{l^2} \quad (7.1)$$

Here μ denotes the strength of the dipole. The meaning of φ and l is

AD-A104 164

OHIO STATE UNIV COLUMBUS DEPT OF GEOMETRIC SCIENCE
THE USE OF FINITE ELEMENTS IN PHYSICAL GEODESY (U)

P/G B/D

APR 81 P MEISSL

F19628-79-C-0075

UNCLASSIFIED

D05-313

AFGL-TR-81-0114

NL

3 1/2 3
0/00 0/04



END
DATE
FILMED
40-81
DTIC

seen from figure 7.1

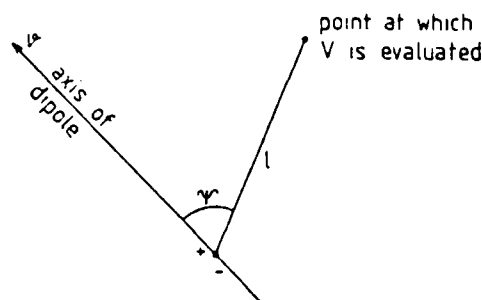


Figure 7.1. Explains the notation of quantities related to a dipole

It is useful to remember that the potential of a dipole is obtained from that one of a mass point by taking the directional derivative with respect to the axis of the dipole and at the location of it.

$$\frac{\cos \psi}{l^2} = \frac{\partial}{\partial \nu} \frac{1}{l} \quad (7.2)$$

It follows that the potential of a dipole decreases like $O\{\frac{1}{r^2}\}$ as $r \rightarrow \infty$. Hence it is free of zero and first order harmonics. A dipole layer is obtained by locating dipoles at a surface and letting the strength of the dipole be a function of location. The axis may also vary with location. Physically and mathematically most meaningful is the coincidence of the axis with the surface normal. The potential of a dipole layer is also $O\{\frac{1}{r^2}\}$.

Dipoles may be generalized to multipoles. Multipoles were already studied by Gauss and Maxwell. An $N+1$ pole of unit strength is obtained as

$$\frac{\partial}{\partial \nu_1} \frac{\partial}{\partial \nu_2} \dots \frac{\partial}{\partial \nu_N} \frac{1}{l} \quad (7.3)$$

Confer e.g. Lense (1954), p. 80 ff. Trying most simple things first, one puts $\nu_1 = \nu_2 = \dots = \nu_N = \nu$, thus letting the N axes coincide and obtaining

$$\frac{\partial^N}{\partial \nu^N} \frac{1}{l} \quad (7.4)$$

The generated potential is $O\left\{\frac{1}{r^{N+1}}\right\}$, i.e. it is free of spherical harmonics up to and including degree $N-1$. The same holds for the potential of a multipole layer. The axes may be chosen in coincidence with the surface normal.

What is the advantage of taking a multipole layer instead of a single layer? At a first glance it appears that multipole layer also generates a full normal equation matrix. Mind, however, that the kernel of a multipole of order $N+1$ decreases as $O\left\{\frac{1}{r^{N+1}}\right\}$ as $r \rightarrow \infty$. This implies that even for moderately large N the contribution of the multipoles located at a small surface element decreases rapidly as one moves away from that surface element. If the surface on which the multipole layer is assumed is subdivided into finite elements, and if the strength function is represented in terms of nodal parameters, then the nodal parameters of elements at a larger distance will be coupled only by very small coefficients which may be replaced by zeroes. Hence a sparse system of normal equations will be obtained.

What are the difficulties to be expected in implementing the multipole layer method, and what is an appropriate finite element representation of the strength function? How smooth should it be chosen? In order to give a (preliminary) answer to these questions we shall take a closer look at spherical multipole layers in the next subsection.

7.3. Multipole layers on the sphere.

Assume a sphere of radius $R = 1$ and let $r > R = 1$. Starting from the well known representation for the function

$$\frac{1}{l} \quad \text{with} \quad l = \sqrt{R^2 + r^2 - 2 R r \cos \psi} \quad (7.5)$$

namely,

$$\frac{1}{l} = \frac{1}{r} \sum_{n=0}^{\infty} P_n(\cos \psi) \left\{ \frac{R}{r} \right\}^n ; \quad r > R \quad (7.6)$$

differentiating N times with respect to R and putting afterwards $R = 1$, one gets

$$\frac{\partial^N}{\partial R^N} \frac{1}{l} = \sum_{n=N}^{\infty} P_n(\cos \psi) n(n-1) \dots (n-N+1) \frac{1}{r^{n+1}} \quad (7.7)$$

Here $P_n(t)$ are the familiar Legendre polynomials, normalized in a way that $P_n(1) = 1$. The series (7.7) converges for $r > 1$. Letting $r \rightarrow 1$, we obtain in the limit a sequence which converges distributionally. The potential generated by the multipole layer with strength function μ is

$$V(r\xi) = \int_{\Gamma} \frac{\partial^N}{\partial R^N} \frac{1}{l(r\xi, R\eta)} \Big|_{R=1} \mu(\eta) d\Gamma(\eta) \quad (7.8)$$

$$r > R = 1$$

Here ξ, η are unit vectors, Γ is the surface of the unit sphere. We temporarily fix r and view $V(r\xi)$ as a function of ξ only. Then (7.8) represents an isotropic operator. According to the theory outlined in Meissl (1971), chapter 3, we know that its eigenfunctions are the spherical harmonics H_{nm} . We can calculate the eigenvalues by the Funk-Hecke formula obtaining

$$\lambda_n \begin{cases} = 0 & ; n < N \\ = \frac{4\pi}{2n+1} & ; n \geq N = 0 \\ = \frac{4\pi}{2n+1} n(n-1) \dots (n-N+1) \frac{1}{r^{n+1}} & ; n \geq N > 0 \end{cases} \quad (7.9)$$

Hence if $\mu(\xi)$ is represented as

$$\mu(\xi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \mu_{nm} H_{nm}(\xi) \quad (7.10)$$

where we now take the surface-spherical harmonics $H_{nm}(\eta)$ as fully normalized, i.e.

$$\frac{1}{4\pi} \int_{\Gamma} H_{nm}^2(\eta) d\Gamma(\eta) = 1 \quad (7.11)$$

we get the representation for $V(r\xi)$ as

$$V(r\xi) = \sum_{n=N}^{\infty} \lambda_n \sum_{m=-n}^{+n} \mu_{nm} \frac{H_{nm}(\xi)}{r^{n+1}} \quad (7.12)$$

We see that $V(r\xi) = O\left\{\frac{1}{r^{N+1}}\right\}$, as announced earlier. Further we note that the operator is singular. The coefficients μ_{nm} for $n < N$ do not contribute to $V(r\xi)$.

If we like to continue $V(r\xi)$ downward to $r = R = 1$, we must impose restrictions upon the μ_{nm} . The following series must converge

$$\sum_{n=N}^{\infty} \sum_{m=-n}^{+n} (\lambda_n \mu_{nm})^2 < \infty \quad (7.13)$$

to ensure that $V(\xi)$ is a member of H_r , the Hilbert space of squared integrable functions defined on Γ . The operator transforming $\mu(\xi)$ into $V(\xi)$ has eigenvalues

$$\lambda_n = \frac{4\pi}{2n+1} n(n-1) \dots (n-N+1) \quad (7.14)$$

It amplifies the high frequencies more than the lower ones. It may be viewed as an operator from the Hilbert space $H_r^{(k+N-1)}$ into $H^{(k)}$, $k \geq 0$.

The space $H_r^{(1)}$ comprises those functions

$$f(\xi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} c_{nm} H_{nm}(\xi) \quad (7.15)$$

for which the following series converges.

$$\sum_{n=0}^{\infty} \sum_{m=-n}^{+n} (n! c_{nm})^2 < \infty \quad (7.16)$$

In a least squares approach we like to have at $r = 1$ a potential $V(\xi) \in H_r^{(2)}$. Hence $\mu(\xi) \in H_r^{(N+1)}$ appears appropriate. Therefore the trial functions should be C^N across element boundaries. In a Ritz-type approach $V(\xi) \in H_r^{(1)}$, $\mu(\xi) \in H_r^{(N)}$ and C^{N-1} continuity across element boundaries would be sufficient.

Remark: One could avoid the smoothness requirements on $\mu(\xi)$ by the Bjerhammar-sphere approach. A sphere of radius $1-\varepsilon$ is placed concentrically into the unit sphere. The density μ is assumed on this sphere. We thus consider an operator

$$\mu((1-\varepsilon)\xi) \rightarrow V(\xi) \quad (7.17)$$

whose eigenvalues are

$$\lambda_n = \frac{4\pi}{2n+1} n(n-1) \dots (n-N+1)(1-\varepsilon)^{n-N} \quad (7.18)$$

$$n \geq N$$

However, in the present context we do not propose to shove difficulties under the carpet by burying the multipole layer underneath the surface of computation.

Suppose now, for the moment, that in (7.8) $V(r\xi)$ is specified and that $\mu(\eta)$ is sought. If the equations are discretized in some way, we

expect them to be very ill conditioned due to the differences in size of the eigenvalues λ_n . While the high frequent components of $\mu(\eta)$ appear to be well determined, the low frequent components are not.

In physical geodesy we do not start from a known $V(\xi)$. Instead we have a discrete set of measurements corresponding to functionals of $V(\xi)$. The situation is more difficult but nevertheless similar. The linear system leading to $\mu(\xi)$ will be ill conditioned.

One may adopt the viewpoint that ill conditioning is equivalent to an improper problem formulation, and one may simply abandon the outlined approach. On the other hand, if insight teaches one that the effects of the ill-conditioning on the final result will be negligible, one may find a way around the numerical difficulties. Of course, there is no hope that insight can be replaced by calculus.

At the present time, no definite answer can be given. More time is needed to think about the method and to conduct numerical experiments.

Appendix A.

Computational effort associated with the partial reduction of a profiled system of normal equations.

The material of this section is rather unsophisticated in nature. However even trivial considerations may be cumbersome if their verification is left to the reader. Therefore we quickly state some formulas on the number of operations needed to solve a structured large system of positive definite equations. A more systematic introduction to this problem area is found in Meissl (1980), chapter 6.

We introduce the well known concept of the profile of a symmetric matrix $A = (a_{ij})$. It comprises all elements a_{ij} such that (1) $i \leq j$, and (2) there exists an element $a_{kj} \neq 0$, $k \leq i$. Referring to figure A.1 we introduce the profile function $p(x)$ indicated by a heavy solid line. The heavy line should actually be a step function because we have a discrete number of equations. However we smear out the discontinuities. This is legitimate if we deal with a large system.

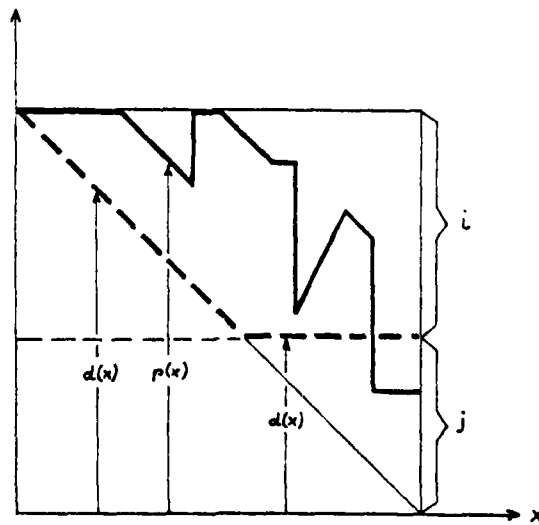


Figure A.1. Partial reduction of a profiled symmetric matrix. Definition of $p(x)$, $d(x)$.

The portion below the main diagonal is not shown in figure A.1. Due to symmetry, there is no need to store these coefficients in the computer.

The upper portion of the system comprises i equations. They are called "interior". The equations below labeled j are called "junction equations". During partial reduction, the interior equations are eliminated. One obtains a system of j partially reduced junction equations.

If the matrix is split as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad (A.1)$$

then the matrix of the partially reduced set is given by

$$A_{22}^{(p)} = A_{22} - A_{21} A_{11}^{-1} A_{12} \quad (A.2)$$

A direct elimination procedure is used such as Gauss, Cholesky or one of the variants. The partial triangular decomposition phase is by far most time consuming. During this phase zero coefficients are enforced below the main diagonal positions of the interior equations. Focus attention on a row of coefficients whose diagonal position is implied by column x . Call this the row x . Let y be a column to the right of x . From the coefficient in row x and column y as many multiples of coefficients located above it are subtracted as the following expression indicates:

$$\Gamma(x, y) = \text{Min} \{ \text{Max}(p(x) - d(x), 0), \text{Max}(p(y) - d(x), 0) \} \quad (A.3)$$

Thereby $d(x)$ is the function implied by the heavy broken line in figure A.1. Hence partial triangular decomposition requires

$$\Gamma = \int_0^{i+j} dx \int_x^{i+j} \Gamma(x, y) dy \quad (A.4)$$

steps. One step comprises one multiplication and one addition. There are also divisions and square roots (in case of Cholesky), but their number is negligible. Also the number of operational steps involving the right hand side, as well as the number of steps arising during the later back-substitution phase are negligible.

Take a fully occupied system. We find

$$\Gamma = \int_0^i dx \int_x^{i+j} x dy + \int_i^{i+j} dx \int_x^{i+j} i dy = \frac{i^3}{6} + \frac{ij(i+j)}{2} \quad (\text{A.5})$$

This formula is frequently used in chapter 4. For $j = 0$, i.e. full reduction, we obtain

$$\Gamma = \frac{i^3}{6} \quad (\text{A.6})$$

References

- Ahlberg J. H.; E. N. Nilson; J. L. Walsh (1967): The theory of splines and their applications. Academic Press, XI + 284 pp.
- Arnold V. I. (1980): Mathematical methods of classical mechanics. Graduate texts in Mathematics, No. 60, second corrected printing. Springer Verlag. X + 462 pp.
- Avila J.; B. Malloy; J. Tomlin (1978): Use of the ILLIAC IV for the readjustment of the North American Datum. Technical Memorandum No. 5732, Institute for Advanced Computation, Sunnyvale, Calif. 94085, 87 pp.
- Bitmead R. R.; B. P. O. Anderson (1980): Asymptotically fast solutions of Toeplitz and related systems of linear equations. Linear Algebra and its applications. Vol. 34, pp. 103 to 116.
- Bjerhammar A. (1964): A new theory of geodetic gravity. Trans. Roy. Inst. Technol., 243, Stockholm.
- Bjerhammar A. (1978): A review of discrete methods in Physical Geodesy. In Moritz-Suenkel (ed.): Approximation methods in Gravity. Wichmann, Karlsruhe, pp. 211 to 233.
- Boehmer K. (1974): Spline-Funktionen. Teubner Studienbuecher, Teubner Stuttgart. 340 pp.
- Bosman E. R.; D. Eckhart; K. Kubik (1972): Delft - a programme system for surface approximation. Bildmessung und Luftbildwesen, Vol. 1972, pp. 13 to 20.
- Brent R. P.; F. G. Gustavson; D. Y. Y. Yun (1980): Fast computation of Pade approximants and the solution of Toeplitz systems of equations. Res. Rep. RC 8173 (#34952), IBM Research Center, Yorktown Heights, NY.

- Ciarlet P. G. (1978): The finite element method for elliptic problems. North Holland Publ. Co., XII + 530 pp.
- Colombo O. L. (1980): Optimal estimation from data regularly sampled on a sphere with applications in geodesy. Air force geophysics laboratory TR-79-0227. Massachusetts.
- Davis T. M.; A. L. Kontis (1970): Spline interpolation algorithms for track-type survey data with applications to mean gravity anomalies. Technical report TR-226, Naval oceanographic office, Washington, D. C. 20390.
- Douglas B. C.; C. C. Goad; F. Morrison (1980): Determination of the geopotential by satellite tracking data. J. of Geophysics. Vol. 85, No. B10, pp. 5471 to 5480.
- Engels R. C.; J. L. Junkins (1979): Local representation of the geopotential by weighted orthonormal polynomials. 17th Aerospace sciences meeting. New Orleans, January 15-17.
- Engels R. C.; J. L. Junkins (1979): Finite element models of the earth's gravity field. Phase IV. Dep. of Engin. Sci. and Mechanics. Virginia Polytechnic Institute and State University. Blacksburg, Virginia.
- Forsberg R.; C. C. Tscherning (1980): The use of height data in gravity field approximation by collocation. Preprint. Geodetic Institute Charlottenlund, Denmark.
- Forsberg R.; F. Madsen (1980): Geoid prediction in northern Greenland using collocation and digital terrain models. Preprint (to be presented at the symposium "Space Geodesy and its Applications", Cannes, November 1980). Geodetic Institute Charlottenlund, Denmark.
- Freedman W. (1978): An application of a summation formula to numerical computation of integrals over the sphere. Bull. Geod., vol. 52, pp. 165 to 175.

Freedden W. (1979): Ueber eine Klasse von Integralformeln der Physikalischen Geodäsie. Veröffentlichungen des Geodätischen Instituts der Rheinisch-Westfälischen Technischen Hochschule Aachen, Nr. 27, 1979. ISSN 0515-0574.

Freedden W. (1980a): Eine Klasse von Kubaturformeln der Einheitssphaere. Manuscript submitted to Zeitschrift fuer Vermessungswesen.

Freedden W. (1980b): Ueber die Gauss'sche Methode zur angenaeherten Berechnung von Integralen. Manuscript submitted to Mathematical methods in the applied sciences.

Freedden W. (1981): Least squares adjustment and smoothing by splines. Manuscript submitted to manuscripta geodetica.

George A. (1973): Nested dissection of a regular finite element mesh. SIAM Journal of Numerical Analysis, 10(2), pp. 345 to 363.

George A.; D. R. McIntyre (1977): On the application of the minimum degree algorithm to finite element systems. Lecture notes in mathematics No. 606, pp. 122 to 149. Springer.

Hardy R. L.; W. M. Goeppfert (1975): Least squares prediction of gravity anomalies, geoidal undulations and deflections of the vertical with multiquadric harmonic functions. Geophysical Research Letters, Vol. 2, No. 10, pp. 423 to 426.

Heiskanen W. A.; H. Moritz (1967): Physical Geodesy. W. H. Freeman and Company. San Francisco and London. 364 pp.

Heller W. G.; K. S. Tait; S. W. Thomas (1977): Geofast - a fast gravimetric estimation algorithm. Air force geophysics laboratory TR-77-0195. Massachusetts.

- Henrici P. (1962): Discrete variable methods in ordinary differential equations. J. Wiley & Sons, XI + 407 pp.
- Jordan S. K. (1978): Fourier physical geodesy. Air force geophysics laboratory TR-78-0056. Massachusetts.
- Junkins J. L. (1977): Development of finite element models for the earth's gravity field. Phase I: Macro gravity model for satellite orbit integration. Rep. No. UVA/525023/ESS77/103. 27 pages.
- Junkins J. L.; R. C. Engels (1979): The finite element approach in gravity modelling. Manuscripta geodaetica Vol. 4, pp. 185 to 206.
- Koch K.; B. Witte (1971): Earth's gravity field represented by a simple layer potential from Doppler tracking of satellites. J. Geophys. Res. Vol. 76(1971), pp. 8471 to 8479.
- Krarup T. (1969): A contribution to the mathematical formulation of Physical Geodesy. Geodetisk Institut, Meddelelse No. 44, Copenhagen, Denmark.
- Leigemann D. (1980): Kollokation und analytische Splinefunktionen. Zeitschrift fuer Vermessungswesen, Vol. 10, pp. 466 to 479.
- Lense J. (1954): Kugelfunktionen. Second edition. Akademische Verlagsgesellschaft Leipzig.
- McDonald B. H.; A. Wexler (1972): Finite element solution of unbounded field problems. IEE Transactions on microwave theory and techniques, Vol. MTT-20, No. 12, pp. 841 to 847.
- McDonald B. H.; M. Friedman; A. Wexler (1974): Variational solution of integral equations. IEE Transactions on microwave theory and techniques, Vol. MTT-22, No. 3, pp. 237 to 248.

- Meissl P. (1969): Ueber zufaellige Fehler in regelmaessigen gestreckten Ketten. Zeitschrift fuer Vermessungswesen. Vol. 94(1969), pp. 14 to 26.
- Meissl P. (1971a): A study of covariance functions related to the earth's disturbing potential. Report No. 151. Dep. of Geodetic science. Ohio State University. IV + 88 pages.
- Meissl P. (1971b): Preparations for the numerical evaluation of second order Molodensky-type formulas. Rep. No. 163. Dep. of Geodetic science. Ohio State University. III + 72 pages.
- Meissl P. (1980): A priori prediction of roundoff error accumulation in the solution of a super-large geodetic normal equation system. NOAA Professional Paper 12. National Geodetic Survey, Rockville, Md. X + 128 pp.
- Moritz H. (1978): Interpolation and approximation. In "Approximation methods in Geodesy" ed. by H. Moritz and H. Suenkel. Sammlung Wichmann, B. 10. H. Wichmann, Karlsruhe.
- Moritz H. (1980): Advanced Physical Geodesy. H. Wichmann, Karlsruhe. XIII + 500 pp.
- Morrison F. (1980): Computing the potential and attraction of a density layer by means of elliptic integrals. Manuscripta geodetica. Vol. 5(1980), pp. 145 to 173.
- Mueller C. (1966): Spherical harmonics. Lecture notes in Mathematics, No. 17, Springer-Verlag, Berlin-Heidelberg-New York.
- Needham P. (1970): The formation and evaluation of detailed geopotential models based on point masses. Rep. No. 149. Dep. of geodetic science. Ohio State University.

- Noë H. (1980): Numerical investigations of the problem of Molodensky. Mitteilungen der geodætischen Institute der Technischen Universitaet Graz, Nr. 30, 80 pp.
- Oden J. T.; J. K. Lee (1977): Dual-mixed hybrid finite element method for second-order elliptic problems. Lecture notes in mathematics No. 606, pp. 275 to 291. Springer.
- Oden J. T. (1979): The classical variational principles of mechanics. In: Glowinsky et al. (ed.): Energy methods in finite element analysis. J. Wiley.
- Pian T. H. H. (1964): Derivation of element stiffness matrices by assumed stress distributions. AIAA-J. 2, pp. 1333 to 1336.
- Rauhala U. A. (1980): Introduction to array algebra. Photogrammetric Engineering and Remote Sensing. Vol. 46, No. 2, pp. 177 to 192.
- Raviart P. A.; J. M. Thomas (1977): A mixed finite element method for 2-nd order elliptic problems. Lecture notes in mathematics, No. 606, pp. 299 to 315. Springer.
- Raviart P. A.; J. M. Thomas (1979): Dual finite element models for second order elliptic problems. In: Glowinsky et al. (ed.): Energy methods in finite element analysis. J. Wiley.
- Robertson W. M. (1978): Spherical geodetic transformations. Rep. No. 1181 of Charles Stark Draper Laboratory. MFT. vol. I and II.
- Schoenberg I. J. (1946): Contributions to the problem of approximation of equidistant data by analytic functions, Quart. Appl. Math. 4, pp. 45 to 99, 112 to 141.
- Schwarz H. R. (1980): Methode der finiten Elemente. B. G. Teubner, Stuttgart. 320 pp.

- Sjoeberg L. (1978): A comparison of Bjerhammar's methods and collocation in physical geodesy. Rep. No. 273 Dep. of geodetic science. Ohio State University.
- Snay R. A. (1978): Applicability of array algebra. Reviews of Geophysics and Space Physics. Vol. 16, No. 3 pp. 459 to 464.
- Strang G.; G. J. Fix (1973): An analysis of the finite element method. Prentice Hall Inc. Englewood Cliffs. XIV + 306 pp.
- Suenkel H. (1977): Die Darstellung geodaetischer Integralformeln durch bikubische Spline-Funktionen. Mitteilungen der geodaetischen Institute der Technischen Universitaet Graz, No. 28.
- Szameitat H. (1979): Modell einer elastischen Erde auf Grundlage finiter Kugelschalenelemente. DGK, Reihe C, Nr. 249. 133 pp.
- Tait K. S. (1980): A fast estimation algorithm for two-dimensional gravity data (Geofast). Air force geophysics laboratory TR-80-0016. Massachusetts.
- Taylor R. L.; O. C. Zienkiewicz (1979): Complementary energy with penalty functions in finite element analysis. In: Glowinsky et al. (ed.): Energy methods in finite element analysis. J. Wiley.
- Thomas S. W.; W. G. Heller (1976): Efficient estimation techniques for integrated gravity data processing. Air force laboratory TR-76-0232. Massachusetts.
- Tong P.; T. H. H. Pian (1969): A variational principle and the convergence of a finite element method based on assumed stress distribution. Int. J. Solids Structures Vol. 5, pp. 463 to 472.

- Tscherning C. C.; R. H. Rapp (1974): Closed covariance expressions for gravity anomalies, geoid undulations and deflections of the vertical implied by anomaly degree variance models. Rep. No. 208 of the Dep. of Geod. Sci. The Ohio State University, Columbus, Ohio.
- Tscherning C. C. (1978): A user's guide to geopotential approximation by stepwise collocation on the RC 4000 computer. Rep. No. 53 of the Danish Geodetic Institute, Copenhagen.
- Tscherning C. C. (1979): Comparison of some methods for the detailed representation of the earth's gravity field. Paper presented at the General Assembly of IUGG/IAG, Canberra, Australia, Dec. 1979.
- Wagner C. A.; O. L. Colombo (1978): Gravitational spectra from direct measurements. Goddard space flight center. Greenbelt, Maryland.
- Werner T. (1979): Untersuchungen zur Erfassung strain-induzierter Störeinflüsse auf Neigungsmessungen. DGK., Reihe C, Nr. 253, 102 pp.
- Zienkiewicz O. C. (1971): The finite element method in engineering science. McGraw-Hill. London.
- Zienkiewicz O. C.; D. W. Kelly; P. Betters (1979): Marriage a la mode - the best of both worlds (finite elements and boundary integrals). In: Glowinsky et al. (ed.): Energy methods in finite element analysis. J. Wiley.
- Zohar S. (1969): The algorithm of W. F. Trench. J. Assoc. Comput. Mach., Vol. 16, pp. 592 to 601.